UNIVERSITY OF CALIFORNIA
RIVERSIDE

Applications of Multi-Cycle Earthquake Simulations to Earthquake Hazard

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Geological Sciences

by

Jacquelyn Joan Gilchrist

December 2015

Dissertation Committee:

Dr. James H. Dieterich, Chairperson
Dr. David D. Oglesby
Dr. Gareth J. Funning

The Dissertation of Jacquelyn Joan Gilchrist is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

Acknowledgments

First of all, thank you to the Southern California Earthquake Center and the National Science Foundation for funding the grants that supported most of this research.

Thank you to my adviser, Jim Dieterich, for all of the years of support and guidance. You have been a brilliant and patient mentor and I am immensely grateful that you have given me the opportunity to work on the subjects that I am most passionate about. It has been an honor to work with and learn from you, and I look forward to continuing to do so.

Thank you to my committee members, David Oglesby and Gareth Funning, for your invaluable support and advice. You provided the perfect balance of brutal honesty and nurturing counsel, while supplying indispensible instruction. You made me a better scientist and a better writer, and helped me learn to believe in myself as well.

Thank you Keith Richards-Dinger for the countless hours you've spent teaching me to understand and use RSQSim, as well as write and debug many of the scripts I used to process the simulated data. I was a poor computer scientist when I arrived at UCR, and I would have been completely lost without your support. Your patience has not gone unappreciated.

Thank you Emily Brodsky for taking a chance on a clueless undergraduate and sharing your passion for geophysics.

Thank you to Jennifer Reising and Laurie Graham for keeping the department running smoothly and putting up with all kinds of desperate maintenance requests and late or incorrectly filed forms. I would have likely frozen to death in my office and never been reimbursed for anything without you both.

Thank you to Eli Brewer, Busra Celikkaya, Laura Bilenker, Sara Beck, Mike Floyd, Julian Lozos, Megan Rohrsson, Erika Noll, Jerlyn Swiatlowski, Kenny Ryan, and Corrie Neighbors for advice, commiseration, and friendship. You helped keep me sane while I navigated the thrilling, but often confusing, and sometimes terrifying waters of grad school.

A special thanks to Eli Brewer and Sara Beck for being wonderful friends and fantastic proofreaders.

To Mom and Don, I wouldn't be here without you. Thank you for supporting me in everything.

To Eli, this would have been so much harder without your unwavering patience and support.

I love you.

ABSTRACT OF THE DISSERTATION

Applications of Multi-Cycle Earthquake Simulations to Earthquake Hazard

by

Jacquelyn Joan Gilchrist

Doctor of Philosophy, Graduate Program in Geological Sciences
University of California, Riverside, December 2015
Dr. James H. Dieterich, Chairperson

This dissertation seeks to contribute to earthquake hazard analyses and forecasting by conducting a detailed study of the processes controlling the occurrence, and particularly the clustering, of large earthquakes, the probabilities of these large events, and the dynamics of their ruptures. We use the multi-cycle earthquake simulator RSQSim to investigate several fundamental aspects of earthquake occurrence in order to improve the understanding of earthquake hazard. RSQSim, a 3D, boundary element code that incorporates rate- and state-friction to simulate earthquakes in fully interacting, complex fault systems has been successful at modeling several aspects of fault slip and earthquake occurrence. Multi-event earthquake models with time-dependent nucleation based on rate- and state-dependent friction, such as RSQSim, provide a viable physics-based method for modeling earthquake processes. These models can provide a better understanding of earthquake hazard by improving our knowledge of earthquake processes and probabilities. RSQSim is fast and efficient, and therefore is able to simulate very

long sequences of earthquakes (from hundreds of thousands to millions of events). This makes RSQSim an ideal instrument for filling in the current gaps in earthquake data, from short and incomplete earthquake catalogs to unrealistic initial conditions used for dynamic rupture models. RSQSim catalogs include foreshocks, aftershocks, and occasional clusters of large earthquakes, the statistics of which are important for the estimation of earthquake probabilities. Additionally, RSQSim finds a *near optimal nucleation location* that enables ruptures to propagate at *minimal stress conditions* and thus can provide suites of heterogeneous initial conditions for dynamic rupture models that produce reduced ground motions compared to models with homogeneous initial stresses and arbitrary forced nucleation locations.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Using Earthquake Simulations to Investigate Earthquake Hazard

This dissertation seeks to contribute to earthquake hazard analyses and forecasting by conducting a detailed study of the processes controlling the occurrence, and particularly the clustering of large earthquakes, the probabilities of these large events, and the dynamics of their ruptures. These studies use the multi-cycle earthquake simulator RSQSim to investigate several fundamental aspects of earthquake occurrence in order to improve the understanding of earthquake hazard. Major topics include: 1. The effect of heterogeneous initial stresses on ground motion models, 2. The effect of different detectability models for paleoseismic events used for probabilistic seismic hazard estimates, 3. The factors that control the timing, location, and rupture extent of large event-clusters, and 4. Whether there are indicators that additional sections of a fault system may rupture shortly after a large event occurs (producing a large event cluster).

The remainder of this chapter explains the significance and advantages of earthquake simulators, as well as the important details about the earthquake simulator RSQSim and how it works. Chapter 2 discusses the characteristics of RSQSim events, and presents an application for generating realistic initial conditions for dynamic rupture models, which are used for ground motion estimates. Chapter 3 details the recently

updated California fault model that is based on the Unified California Earthquake Rupture Forecast Version Three (UCERF3) Fault Model, explains the tuning process for the simulated California catalogs, and explores different models of paleoseismic event detectability in an effort to explain the modern 'hiatus' in large California earthquakes. Chapter 4 focuses on the statistics of the simulated, long-term catalogs with particular emphasis on clusters of large earthquakes. Chapter 5 focuses on the foreshocks and aftershocks of large events in simulated catalogs and how those statistics can be used to estimate the probabilities of large-event clusters. Finally, Chapter 6 presents a brief synthesis of findings and suggests directions for future research.

1.2 Earthquake Hazard and the Necessity of Simulations

Understanding earthquake occurrence, in terms of the timing or statistics of events, as well as the role of fault stresses on recurrence, triggering, and ground motions, is important for earthquake hazard estimation. However, there is not sufficient data on the occurrence of large earthquakes to effectively forecast the timing, location, and intensity of ground motions. Most of the large earthquakes that represent the greatest hazard have recurrence times on the order of hundreds or thousands of years. Instrumental catalogs span just over one hundred years, but the early data are sparse and inexact. Historical and paleoseismic records provide information about the long-term recurrence of events and sometimes the intensity of ground motions in some select locations; but, these records have large uncertainties and lack important details such as well constrained magnitudes of events and the statistics of aftershock sequences. The earthquake simulator RSQSim,

however has been successful in producing realistic earthquake catalogs that reproduce several aspects of earthquake occurrence, including several modes of earthquake slip, and empirical aftershock and magnitude-frequency relationships. The studies described in the following chapters use RSQSim with a California fault model to generate various long-term, simulated catalogs to address many of the questions left unanswered by paleoseismic and instrumental catalogs, as well as simple earthquake rupture models.

1.3 Introduction to RSQSim

RSQSim is a 3D, quasi-dynamic, boundary element model that incorporates rate- and state-dependent friction to simulate earthquakes in fully interacting, complex fault systems. It is capable of simulating both small (single element, ~M3) and large (10^3 elements, ~M9) seismogenic events, in simple and complex fault systems, in the form of foreshocks, aftershocks, and mainshocks (Richards-Dinger and Dieterich, 2012). Additionally, RSQSim is capable of simulating fault creep and slow-slip events (Colella et al., 2011), as well as injection-induced seismicity (Dieterich et al., 2015).

1.3.1 RSQSim Computations

RSQSim uses the Ruina (1983) and Rice (1983) simplification of the Dieterich (1981) formulation of the rate- and state-constitutive law to calculate the sliding resistance between fault surfaces:

$$\tau^{fric} = \sigma\left[\mu_0 + a\ln\left(\frac{V}{V_0}\right) + b\ln\left(\frac{\theta V_0}{D_C}\right)\right], \tag{1.1}$$

where τ^{fric} is the frictional shear stress that is resisting motion on the fault; σ is the normal stress; a is the rate coefficient; b is the state coefficient; V is the slip speed; V_0 is a reference slip speed; θ is the frictional state variable that evolves with time, slip, and normal stress history; D_C is the characteristic slip distance over which θ evolves; and μ_0 is the steady-state coefficient of friction (at the reference slip speed). Material parameters μ_0, a, b, and D_C are all experimentally measured constants that are specified for each simulation. Unless otherwise specified, the following studies use typical values for the constants μ_0=0.6, a=0.01, b=0.014, and D_C =10^{-5}m (Richards-Dinger and Dieterich, 2012). The state variable has dimensions of time and it evolves with time, displacement, and normal stress as given by Linker and Dieterich (1992) and Dieterich (2007):

$$d\theta = dt - \left(\frac{\theta}{D_c}\right)dV - \left(\frac{\alpha\theta}{b\sigma}\right)d\sigma\,, \tag{1.2}$$

where α is a scale factor corresponding to the slope of the line representing the change in coefficient of friction during sliding versus the amplitude of a normal stress perturbation. θ evolves over a characteristic sliding distance D_c and at steady-state:

$$\theta^{SS} = \frac{D_c}{V}\,. \tag{1.3}$$

Substituting Equation 1.3 into Equation 1.1 gives the steady-state friction as a function of slip speed:

$$\tau^{SS} = \sigma\left[\mu_0 + (b-a)\ln\left(\frac{\theta V_0}{D_C}\right)\right]. \tag{1.4}$$

RSQSim employs a highly efficient computational approach based on Dieterich (1995) that allows it to generate synthetic catalogs on the order of 10^6 M3.5-M8 earthquakes. This is made possible with use of event-driven computational steps of variable length rather than time stepping at closely spaced intervals (Dieterich & Richards-Dinger, 2010; Richards-Dinger & Dieterich, 2012). The cycle of stress accumulation and earthquake slip at each fault segment is separated into three distinct sliding states 0 (locked), 1 (nucleating), and 2 (sliding). Figure 1.1 is a snapshot of a single fault element during a large RSQSim event indicating when that element is either locked (red), nucleating (blue), or sliding (green). This figure also illustrates the evolution of slip, shear stress, normal stress, coefficient of friction, and sliding state for the same fault element and time period. The sequence of events leading up to a large earthquake in RSQSim simulations begins with stress accumulation due to tectonic stressing and stress transfer from slip of nearby fault elements. During this phase the fault is locked. A fault element is in state 0 if the shear stress is below the steady-state friction. The frictional state-variable increases over time (as given by Equation 1.2) and the fault strengthens until the applied shear stress exceeds the steady-state shear stress (Equation 1.4) and the fault element transitions to state 1. During state 1 conditions have not been met for unstable slip, but the fault progressively weakens. The nucleation solutions of Dieterich (1992) modified to account for normal stress changes (Dieterich, 2007), together with stressing rate determine the event transition time to state 2. The time to instability is highly dependent on stressing rate and nucleation often takes years at tectonic stressing rates, but can drop to <1 s at high stressing rates during earthquakes.

Once unstable (seismic) slip begins, the fault element transitions to state 2 and the shear stress and coefficient of friction decrease until the applied shear stress drops below the steady state shear stress, and the element stops slipping (transitions back to state 0). It is not uncommon for individual elements to lock (as stress drops below steady state due to slip on that element) and re-nucleate (as more stress is transferred from slip of other elements nearby) multiple times during a large event as seen in Figure 1.1. Changes in normal stress during rupture are due to fault geometry and are greatest during ruptures on curved or dipping faults. In the case of the fault element in Figure 1.1, the normal stress decreases due to unclamping of the fault as the rupture propagates along the curved surface. In state 1, slip is accelerating but the total slip is negligible (a few times D_c) and therefore ignored. However, in state 2, a fault element has slipped sufficiently far to be transferring stress to every other element in the model.

The fault models employed in this study use either rectangular 3D boundary elements based on Okada (1992) or triangular elements based on Gimbutas (2012). Boundary element formulations for RSQSim are represented by 3D elastic dislocation arrays based on interactions between fault elements. Shear and normal stresses acting on elements are:

$$\tau_i = K_{ij}^{\tau}\delta_j + \tau_i^{tect} \tag{1.5}$$

$$\sigma_i = K_{ij}^{\sigma}\delta_j + \sigma_i^{tect} \tag{1.6}$$

Where i and j are from 1 to N (the total number of fault elements), τ_i is the shear stress in the direction of slip on the i^{th} element, and σ_i is the fault-normal stress on the same element. The two K_{ij} are interaction matrices derived from either the Okada (1992) or

Gimbutas (2012) dislocation solutions; slip on fault element j is represented by δ_j and tectonic stresses at fault element i are denoted by τ_i^{tect} and σ_i^{tect}. The interaction (stiffness) matrices shown in Equations 1.5 and 1.6 give the stress change on the centers of each element due to unit slip on any other element. The stiffness matrices and the tectonic stressing rates are calculated once at the start of each simulation to save computation time. The long-term slip rate is specified for each element in the model and the back-slip loading method is used to set the tectonic stressing rates. With this method the fault elements are forced to slip backwards at their long-term rate and the stiffness matrix is used to solve for the rate of stress change on each element. These back-slip stressing rates are then used to load the fault and are exactly negated over the long-term when faults slide in the forward direction at their long-term rates.

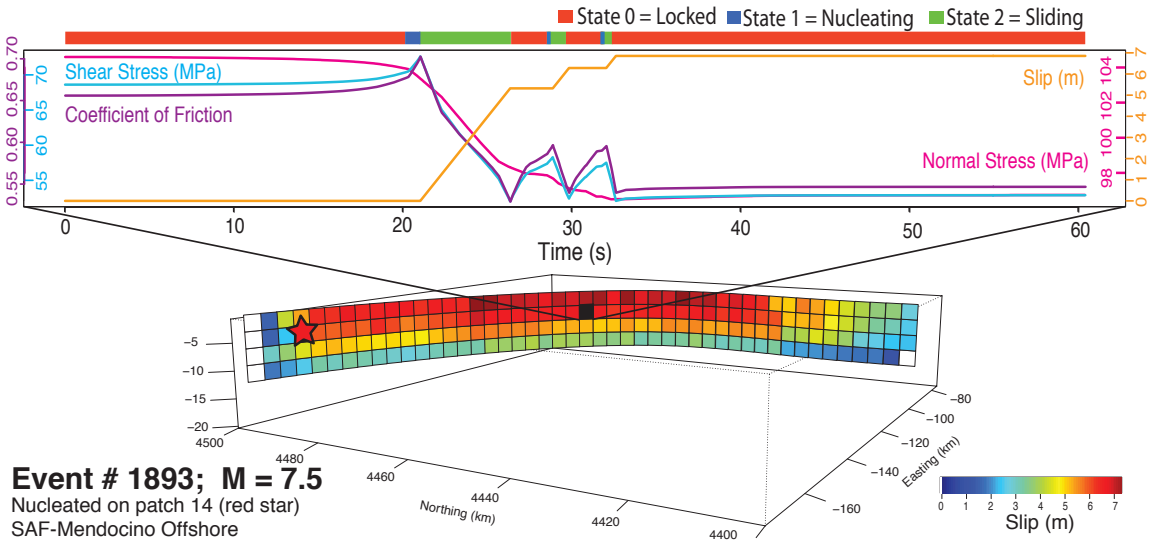

Figure 1.1: Evolution of slip (orange line), shear stress (blue line), normal stress (pink line) and the coefficient of friction (purple line) with time (upper panel) for a single fault element during a single large event (the full slip distribution of the event is shown below) from RSQSim. The sliding state for this element is indicated by the colored bar: red = State 0 = locked, blue = State 1 = nucleating, and green = State 2 = sliding. Colors on the fault image (lower panel) indicate total cumulative slip for all elements during this event. The red star indicates the hypocenter and the black patch indicates the fault element for which the history is plotted.

1.3.2 California Fault Models

The initial fault model used in this study is a California fault model based on the UCERF2 fault model (Figure 1.2). It is a version of the Allcal2 fault model described by Tullis et al. (2012), modified to run in RSQSim. It includes most of the mapped faults in California and a few that extend into Northern Mexico and run offshore of the Pacific coast. Back-slip tectonic stressing rates for each fault are based on the geologic fault slip rates from the UCERF2 report. RSQSim can accommodate any mode of faulting i.e. normal, reverse, strike-slip, and oblique, and modeled faults can be in any orientation. However, the rake angle for each element is set *a priori* and remains fixed. An approximation built into this fault model is that slip rate tapers to zero over a few patches both down-dip and at the ends of the faults in order to avoid sharp boundaries in stress during simulations. In total, the fault model contains 12,500 elements, at a 3x3 km resolution. Since earthquake moment, and therefore earthquake magnitude, depends on fault area this resolution corresponds to a minimum possible earthquake magnitude of M4.9 for an event on a single element. The simulated catalogs produced with this model each span tens of thousands of years and contain a few hundred thousand M4.9 to M8 events. It should be noted that there is some 'run-up' time for each simulation that depends on the minimum slip-rate in the model. For this reason, the beginning of each simulated catalog (typically ~1,000 years) is thrown out to avoid irregularities in the first ruptures on each fault.

While preliminary results from the All California model provide insight into earthquake clustering, longer catalogs with a much wider magnitude range are required

for better clustering and aftershock statistics. An updated California fault model, based on the fault model and slip rates from the UCERF3 report was developed in 2014. This updated (UCERF3) fault model, which is discussed in more detail in Chapter 3, has more than a quarter of a million triangular elements at a resolution of 1 km^2. This reduces the minimum magnitude to roughly magnitude 3.5, and greatly improves the aftershock statistics, as detailed in Chapter 5.

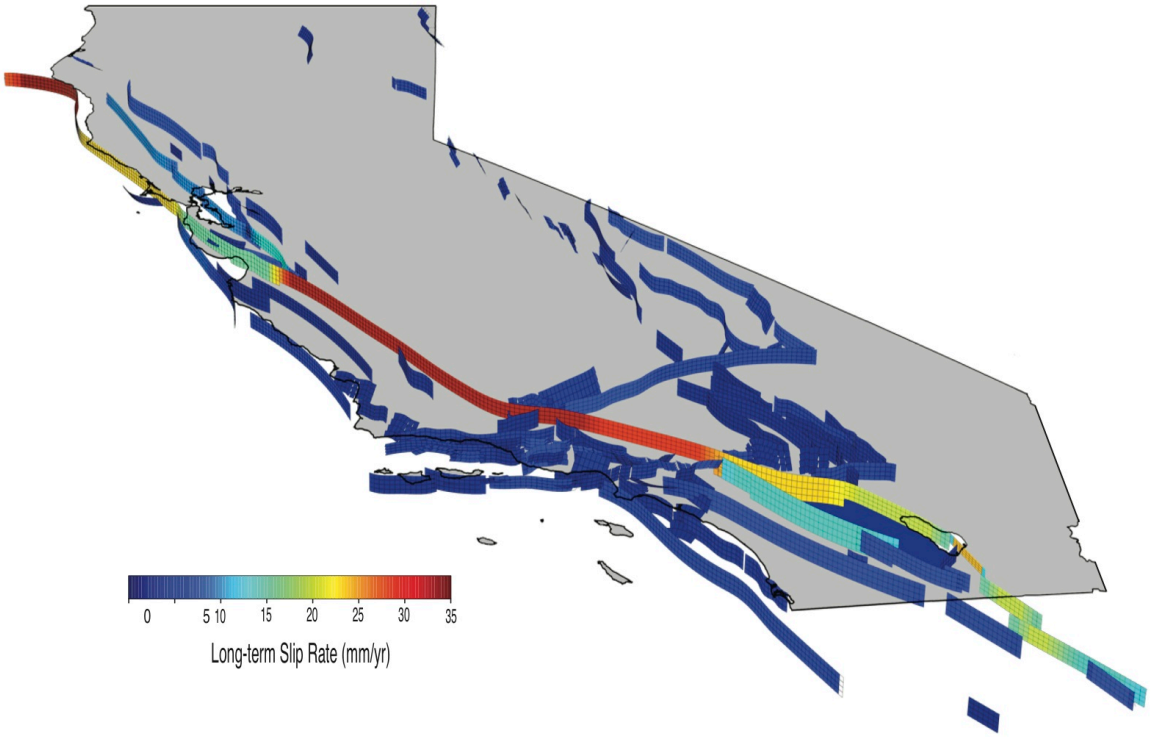

Figure 1.2: All California Fault Model. Element sizes are ~3km by ~3km with some variation to match mapped faults as well as possible. Color indicates the long-term slip-rate built into the model, based on observations reported by the UCERF2 report.

1.3.3 Evaluating RSQSim for Studying Long-Term Earthquake Behavior

Use of rate- and state-friction results in space-time clustering that follows the Omori aftershock decay law and permits investigation of earthquake clustering (Richards-Dinger and Dieterich, 2012). RSQSim allows large-scale simulations of fault system seismicity, which produce clustered seismicity in the form of foreshocks, aftershocks, and occasional large-event clusters. As shown in Chapters 2 and 3, RSQSim generates realistic earthquake catalogs that are long enough to provide useful statistics. It creates synthetic catalogues that span hundreds of thousands of years, with hundreds of thousands to a million events. While the catalogs span up to a million years, we do not claim to make any predictions about fault slip rates or geometry in the future. The long catalogs are necessary for statistical analysis of events and long-term earthquake probability estimates, particularly for events on faults with slow slip rates

Some basic catalog statistics from a representative simulation with the UCERF2 fault model are shown in Figure 1.3. The fence plot of earthquake magnitudes with time (Figure 1.3a) shows great variation in the occurrence of large events, including periods of partial quiescence, as well as an increase in both the number and magnitude of smaller events surrounding the occurrence times of large events. The magnitude-frequency plot (Figure 1.3b) roughly follows Gutenberg-Richter scaling with a typical b-value around 1. The inter-event waiting time distribution (Figure 1.3c) shows the probability density of the times between events in the catalog. This plot has a power-law slope that rolls off to approximately follow an exponential distribution (shown by the blue curve). The exponential curve gives the probability density for a Poisson distribution. The

magnitude–rupture area plot (Figure 1.3d) approximately fits the Wells and Coppersmith (1994) scaling relationship.

The aftershocks in the RSQSim catalogs show Omori-like decay. However, the widely observed strong dependence of aftershock productivity on earthquake magnitude (Shearer, 2012) is not seen because off-fault seismicity is not modeled. Earthquakes can only occur on faults that are explicitly modeled, and consequently the aftershocks are largely limited to ends of mainshock ruptures. Additionally, the minimum rupture area and thus magnitude of events is controlled by the element size of the model.

Mid-scale simulations of $\sim 10^5$ events in a fault system with approximately 2×10^4 fault elements, such as the one shown in Figure 1.2, can be run on desktop computers in a few hours. RSQSim has the ability to restart simulations with the ending conditions of previous simulations, allowing extensions to the timescale and number of events of any catalog. In this way multiple, continuous catalogs have been combined to create catalogs that span up to a million years with tens of millions of events. Midway through this study, RSQSim was parallelized, permitting larger simulations that run on supercomputers or local clusters. Longer and higher resolution simulations were run on the Linux cluster in the UCR Earthquake Physics Laboratory and the highest resolution models were run on the Stampede Supercomputer at the Texas Advanced Computing Center in Austin, Texas.

In summary, multi-event earthquake simulators with time-dependent nucleation based on rate-state friction, such as RSQSim, provide a viable physics-based method for modeling earthquake processes. RSQSim catalogs compare well to California catalogs for the magnitude ranges that are available for comparison. Richards-Dinger and

Dieterich (2012) participated in a simulator comparison project (Tullis et al., 2012b) that determined that RSQSim was in good agreement with California data, and performed as well or better than the other three simulators tested in the study. Additionally, Richards-Dinger and Dieterich (2012) determined that RSQSim ruptures compare well to ruptures from fully dynamic rupture models in terms of total slip, rupture time history, and stress change. Overall, given the variety of fault models RSQSim can accept, the length of catalogs it can produce, and success of modeling different aspects of fault slip and earthquake occurrence, RSQSim is very well suited for studying both short- and long-term earthquake processes and statistics.

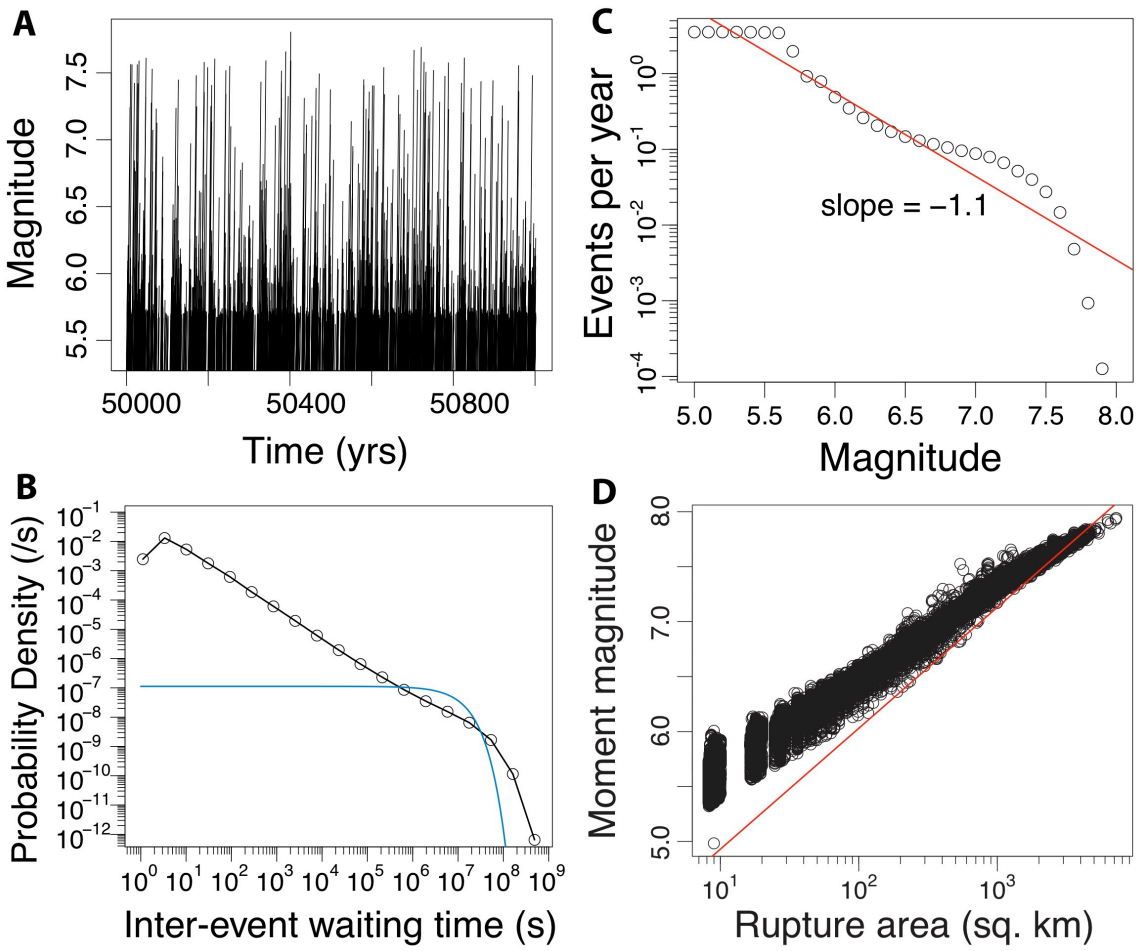

Figure 1.3: a. Magnitude frequency distribution for a 1000 year period in the middle of a long-term catalog, b. Probability density of Inter-event waiting times with the exponential fit to the data shown in blue, c. Gutenberg-Richter plot with the fit line for the slope shown in red, and d. Scaling relationship for moment magnitude with rupture area, red line is the Wells and Coppersmith (1994) moment-area relationship for all rupture types.

1.4 References

Colella, H. V., Dieterich, J. H., & Richards‐Dinger, K. B. (2011). Multi‐event simulations of slow slip events for a Cascadia‐like subduction zone. *Geophysical Research Letters*, *38*(16).

Dieterich, J. H., (1992) Earthquake nucleation on faults with rate and state-dependent friction, Tectonophysics, 211, 115–134.

Dieterich JH (1981) Constitutive properties of faults with simulated gouge. In: Carter NL, Friedman M, Logan JM, and Stearns DW (eds.) Monograph 24: Mechanical Behavior of Crustal Rocks, pp. 103–120. Washington, DC: American Geophysical Union.

Dieterich, J. H. (1995). Earthquake simulations with time-dependent nucleation and long-range interactions. *Nonlinear Processes in Geophysics*, *2*(3/4), 109-120.

Dieterich, J. (2007). Applications of rate-and-state-dependent friction to models of fault slip and earthquake occurrence, in Treatise On Geophysics, G. Schubert (Editor), Vol. 4, Elsevier, Oxford.

Dieterich, J. H., and K. Richards-Dinger (2010), Earthquake recurrence in simulated fault systems, Pure Appl. Geophysics., 167,1087–1104. doi:10.1007/s00024-010-0094-0.

Dieterich, J. H., Richards‐Dinger, K. B., & Kroll, K. A. (2015). Modeling injection‐induced seismicity with the physics‐based earthquake simulator RSQSim. *Seismological Research Letters*.

Gimbutas, Z., Greengard, L., Barall, M., & Tullis, T. E. (2012). On the calculation of displacement, stress, and strain induced by triangular dislocations. *Bulletin of the Seismological Society of America*, *102*(6), 2776-2780.

Linker, M. F., and J. H. Dieterich (1992). Effects of variable normal stress on rock friction—Observations and constitutive equations, J. Geophysics. Res. 97, 4923–4940.

Okada,Y. (1992). Internal deformation due to shear and tensile faults in a half-space, Bull. Seismol. Soc. Am. 82, 1018–1040.

Rice, J. R. (1983). Constitutive relations for fault slip and earthquake instabilities. *Pure and applied geophysics*, *121*(3), 443-475.

Rice, J. (1993). Spatio-temporal complexity of slip on a fault. *Journal of Geophysical Research*, *98*(B6), 9885-9907.

Ruina, A. (1983). Slip instability and state variable friction laws. *J. geophys. Res*, *88*(10), 359-10.

Richards-Dinger, K., and J. H. Dieterich (2012), RSQSim Earthquake Simulator, Seismological Research Letters, v. 83, no. 6, p. 983-990, 2012.

Savage, J. C. (1983). A dislocation model of strain accumulation and release at a subduction zone. *J. geophys. Res*, *88*(6), 4984-4996.

Shearer, P. M. (2012). Self-similar earthquake triggering, Båth's law, and foreshock/aftershock magnitudes: Simulations, theory, and results for southern California. *Journal of Geophysical Research: Solid Earth (1978–2012)*, *117*(B6).

Tullis, T. E., K. Richards-Dinger, M. Barall, J. H. Dieterich, E. H. Field, E. M. Heien, L.H. Kellogg, F. F. Pollitz, J. B. Rundle, M. K. Sachs, D. L. Turcotte, S. N. Ward, and M. B. Yikilmaz (2012b). Comparison among observations and earthquake simulator results for the allcal2 California fault model, Seismol. Res. Lett. 83, 994–1006.

Wells, D. L., & Coppersmith, K. J. (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bulletin of the seismological Society of America*, *84*(4), 974-1002.

Chapter 2

Characteristics of RSQSim Events and the Importance of Evolved Stress States

2.1 Abstract

Initial conditions used for dynamic rupture models, particularly the pattern of initial stresses, play a primary role in controlling rupture characteristics and resulting ground motions. The earthquake simulator RSQSim is able to simulate very long sequences of earthquakes (typically >100,000 events) and thus generate suites of heterogeneous initial conditions for large rupture events. In these simulations, earthquakes nucleate spontaneously — hence, each observed rupture has a set of initial stress conditions and an associated nucleation location that is not set *a priori* or forced. One important aspect of the simulations, and presumably of natural fault systems, is that repeated small events (prior to a large, through-going event) test the ability of a rupture to propagate through the system, and by this process the system finds a *near optimal nucleation location* that enables ruptures to propagate at a *minimal stress condition*. Forcing nucleation at non-optimal locations for a given stress condition requires system-wide increases in shear stress to produce through-going ruptures. Compared to ruptures with homogeneous initial stresses and arbitrary forced nucleation locations, the spontaneous RSQSim events are able to propagate at lower average initial stresses and thus have lower stress drops (by

roughly 50%) and lower rupture propagation speeds, which produces significantly lower ground motions.

2.2 Introduction

Ground motion estimates are a key aspect of earthquake hazard maps and provide the primary information for designing earthquake resistant structures. Near-field ground motion estimates are essential for hazard estimation and disaster planning, including zoning, building codes, and post-earthquake relief efforts. However, there are limited data on near-field ground motions from large earthquakes. The largest events that represent the greatest hazard generally have longer recurrence intervals, so there are few in the instrumental record. Initial stress conditions and earthquake ground motions show great variability, possibly due to local factors such as geology and fault geometry, so the data available are not necessarily applicable to regions outside the area in which they were recorded. Many earthquake scientists have turned to dynamic rupture modeling to estimate ground motions in order to characterize the largest and rarest events. Most ground motion calculations so far have been performed with kinematic models; however, increases in computational efficiency and access to more powerful computers has led several earthquake scientists toward the use of physics-based dynamic rupture models for estimating ground motions, because they are now able to model the shaking frequencies (>1 Hz) necessary for the engineering of critical structures (Cui et al., 2010). Other studies have taken a hybrid approach by combining kinematic, stochastic, and dynamic

rupture models to produce broadband ground motions (Mai & Olsen, 2010; Mena et al. 2010; and Lozos et al. 2015).

Initial conditions used for dynamic models, particularly the heterogeneity of initial stresses, play a fundamental role in determining rupture characteristics and resulting ground motions (Andrews, 1980; Day, 1982; Beroza and Mikumo, 1996; Oglesby and Day, 2002; Duan and Oglesby, 2006; Duan and Oglesby, 2007; Lapusta and Liu, 2009). In natural fault systems, initial stress conditions are a product of prior slip history and fault-system interactions. Large-scale fault geometry as well as small-scale roughness affects the stress patterns on faults (Okubo and Dieterich, 1984; Duan and Oglesby, 2007; Candella et al., 2009; Dieterich & Smith, 2009) and it has been established that faults are rough and that their geometry varies at all scales (Power et al., 1987; Sagy and Brodsky, 2009). Furthermore, the degree of stress heterogeneity, for both shear and normal stress, and rupture complexity (i.e. fluctuations in speed, shape, and propagation direction) affects the magnitude and spectrum of ground motions (Madariaga, 1977; Hartzell and Archuleta, 1979; Olsen et al., 2008).

Current methods of specifying initial stresses for modeling dynamic rupture and estimating ground motions fall into 3 categories: homogeneous methods that use uniform initial stresses, stochastic methods that use random or idealized stress patterns (Andrews, 1980; Day, 1982; Oglesby and Day, 2002; Andrews and Barall, 2011), and inversion methods that use waveform inversions to estimate the initial stress field of specific significant earthquakes (Beroza and Mikumo, 1996; Olsen et al., 1997). Some methods of specifying heterogeneous initial stresses have used discrete patches of high and low stress

(Day, 1982; Lapusta and Liu, 2009), but many heterogeneous stress patterns are represented by stochastic distributions of stresses (Andrews, 1980; Oglesby and Day, 2002; Andrews and Barall, 2011). Day (1982) compared dynamic models with homogeneous initial stresses to models with heterogeneous initial stresses and found that even the simplest heterogeneous initial stress patterns (i.e. elements of high shear stress surrounded by lower shear stress) resulted in lower average rupture velocities (compared to models with homogeneous shear stress) with local supershear rupture velocities, where the rupture velocity is higher than the shear wave velocity, in areas with the highest shear stress. This indicates that stress heterogeneities can have an important effect on the dynamics of earthquake rupture, both in nature and in models, and thus on ground motion estimates as well. More complicated stochastic stress patterns attempt to parameterize small-scale heterogeneities in fault geometry that arise from fault roughness and small, nearby faults. The advantage of these stochastic stress patterns is that they account for the unknown and potentially unknowable small-scale structure of faults. However, it is impossible to determine whether these stress distributions are representative of natural stresses, and idealized stress distributions are not customized to any specific fault or system. Furthermore, these methods have no way of coupling the nucleation location to the stress distribution. Inversion methods for specifying heterogeneous initial stress from source inversions are arguably more realistic than stochastic models because they are based on real world data for a specific fault and event. However there is a small sample of ground motion data with which to perform these inversions, the results of which are often non-unique and highly variable because of the numerous, non-unique assumptions

that go into those inversions (Mai 2007; Mai et al. 2007) and the inversions cannot resolve fine-scale, heterogeneous stresses of interest for ground motion simulations. Additionally, these heterogeneous stresses only apply to a single, specific event and do not provide a general rule for assigning initial stresses. The earthquake simulator RSQSim has the ability to generate suites of heterogeneous initial stress conditions that spontaneously arise from long-term loading and fault interactions (Dieterich and Richards-Dinger, 2010) and, we argue, these evolved stresses provide more realistic representations of the stress states of faults immediately prior to large earthquakes, even for the idealized fault geometry used in this study.

2.3 Methods and Models

Two different, yet complementary physics-based methods were implemented to study the effect of initial stresses on earthquake ground motions. The following study, which uses the long-term earthquake simulator RSQSim, and the dynamic earthquake rupture code FaultMod (Barall, 2008), investigates both the evolution of fault stresses over many earthquake cycles and the resulting ground motions from earthquakes that occur under those evolved stress conditions. RSQSim is used to study the characteristics of earthquakes, specifically the magnitudes and variability of initial stresses prior to large events, while FaultMod is used to study the strength and variability of ground motions due to the variations in the evolved stress states from RSQSim events.

2.3.1 RSQSim

As described in Chapter 1, RSQSim (Rate-State earthQuake Simulator) is a 3D boundary-element code developed by James Dieterich and Keith Richards-Dinger. RSQSim incorporates rate-state fault friction to simulate long sequences of earthquakes in complex, fully interacting fault systems (Dieterich and Richards-Dinger, 2010; Richards-Dinger and Dieterich, 2012). It creates synthetic catalogs that span tens of thousands of years, with hundreds of thousands of roughly magnitude 4 to magnitude 8 events. RSQSim is able to handle fault models that are divided into thousands to half-a-million fault elements. During seismic rupture, every fault element that slips transfers stress to the rest of the model; however, there are no dynamic (i.e. seismic wave) interactions. These fault interaction stresses are superimposed upon the long-term tectonic stressing rate that is applied to every fault element in the model. The nucleation location for each individual earthquake is not specified, as events nucleate spontaneously wherever the conditions for nucleation occur (as determined by the nucleation solutions of Dieterich, 1992). RSQSim is able to incorporate either rectangular or triangular fault elements, and it is possible to perform simulations with simple, single planar faults or complicated fault models with varying degrees of curvature and roughness, as well as multiple faults and different fault types.

RSQSim is capable of modeling a wide range of temporal and spatial scales because the code employs a highly efficient computational approach. Unlike dynamic models, which require time stepping at closely spaced intervals, RSQSim uses event-driven computational steps, which make it significantly faster than dynamic models. As

explained in Chapter 1, the length of each time step is variable and determined by calculating the nearest time at which any element in the model will begin to nucleate. The simulator then jumps to that time and adjusts the stresses in the model accordingly. This is significantly faster than time stepping at small, even intervals. In fact, several hundred-thousand RSQSim events can be simulated in the time it takes to perform one single-event, dynamic model with FaultMod.

2.3.2 FaultMod

FaultMod is a 3D, dynamic, finite element code designed to model earthquake fault rupture, slip, wave-propagation, and ground motion. The code was developed by Michael Barall of Invisible Software (Barall, 2008) for physics-based modeling of 3D fault systems. It has been validated by the Southern California Earthquake Center's Code Validation Project and accepted to be a fast, efficient, and most importantly, accurate code (Harris and Archuleta, 2004; Harris et al. 2009). It allows for moderately complex fault geometry (i.e. several faults with bends, stepovers, etc., but nothing as large or complex as all of California, due to gridding limitations), heterogeneous initial parameters, as well as 3D variation of material properties throughout the model. It incorporates several fault constitutive models including linear slip weakening, rate- and state-dependent friction, and Coulomb friction (Barall, 2008). However this study only presents results from the rate- and state-dependent friction model to correspond with RSQSim simulations. Unlike RSQSim, FaultMod requires forced nucleation to initiate dynamic rupture and nucleation is achieved by either reducing the value of the state

variable or increasing the shear stress within an expanding, circular nucleation region such that the slip rates in this region increase to coseismic slip rates. This study nucleates events by reducing the state variable within the nucleation region, which essentially weakens the fault within that region. This ensures that nucleation does not generate excessive rupture speeds or additional slip in the regions of the fault near the nucleation location as is often observed in models nucleated by increasing the shear stress in the nucleation region. The minimum nucleation radius for unstable slip is determined by Equation 15 in Dieterich (1992). However, depending on the level and pattern of initial stresses, rupture will often stall when the minimum nucleation radius is used, therefore we use nucleation diameters that are 10-20% higher than the minimum to ensure that rupture continues to propagate. The diameter of these enlarged nucleation regions is usually between 2 km and 4 km.

Identical values of frictional parameters a, b, and μ_0 (from Equation 1.1) were used in both codes. However, the codes require very different values (by ~3 orders of magnitude) for the rate- and state-dependent friction parameter D_C. RSQSim assumes nucleation occurs within a single element and uses laboratory values of D_C (on the order of 10 μm) which are required to nucleate small earthquakes. However, FaultMod requires large (a few cm) values of D_C in order to properly resolve the evolution of the state parameter within computationally feasible limits on size of the finite element grid. This necessitates tuning the other input parameters to produce the same results for slip, stress drop, and rupture velocity from both RSQSim and FaultMod. This is achieved by adjusting the shear stress in order to modulate the stress drop.

2.4 Results

2.4.1 Evolved Stress States from RSQSim

RSQSim simulations begin with homogeneous initial shear stress, but during multi-event simulations, the shear stress evolves through fault system interactions over many earthquake cycles. These initial stresses continuously evolve through the occurrence of many earthquakes of different sizes, slip distributions, and locations, which act to redistribute stresses. Hence, each rupture initiates at a different heterogeneous initial stress state. Stress heterogeneity is also affected by slip on non-planar faults, complex structure of fault systems, and by non-uniform loading, however the examples in this chapter are all taken from simulations with simple faults with uniform long-term stressing rates. Figure 2.1 shows examples of ten large events ($M \geq 7$) from a 200,000-event simulation with a single, planar, strike-slip fault, uniform normal stress, and uniform constitutive properties. The color and the contours represent the time at which each part of the fault first ruptured. The areas where several rupture front contours overlap indicate local pauses of rupture propagation for several seconds. There is great variation in the rupture-time patterns between events in the same simulation. Because model parameters and fault geometry are fixed, this variability is due to the continuing evolution of shear stresses within the simulation. Events that result from evolved, heterogeneous initial stresses in RSQSim are characteristically different from events initiated with homogeneous or composed stresses. Compared to ruptures with uniform initial stresses, RSQSim events that result from evolved initial stresses are able to propagate at lower average initial stresses. An example of this can be seen in Figure 2.2. This figure shows

the total slip, initial shear stress, normal stress, and rupture time contours of two events that were roughly 5,000 years apart in the same multi-event simulation with initially uniform shear stress and non-uniform normal stress (in the form of 'strong' patches of high normal stress). The first event was artificially nucleated by increasing the shear stress by 5 MPa on one element on the right end of the fault. This was done to avoid nucleating near one of the patches of higher normal stress, which requires greater changes in shear stress or fault strength. The evolved event has a lower initial shear stress (by several MPa), as well as a slower (by a factor of 3) and more variable rupture propagation speed, resulting in a maximum slip that is nearly 3m less than that of first event. Overall, this produces more complex ruptures, such as those in Figure 2.1. There is variation in all of these characteristics throughout the simulations, but the greatest difference is evident between the first event of each simulation and the subsequent events. This is illustrated in Figure 2.3, which shows the initial shear stress, stress drop, rupture duration, and rupture velocity for roughly one hundred full fault rupturing ($M \geq 7$) events (plotted against the nucleation time of each event) in an RSQSim simulation (the same one that the events illustrated in Figure 2.2 came from) that was initiated with homogeneous shear stress at 75MPa. The initial shear stresses and stress drops are averaged over all fault elements, and the rupture velocity is calculated from the distance between the nucleation point and the last element that began to slip, and the time between nucleation and the time at which the last element began slipping. While the average initial stresses of these evolved events vary by a 1-2 MPa, they never reach the magnitude of the initial shear stress prescribed for the first event.

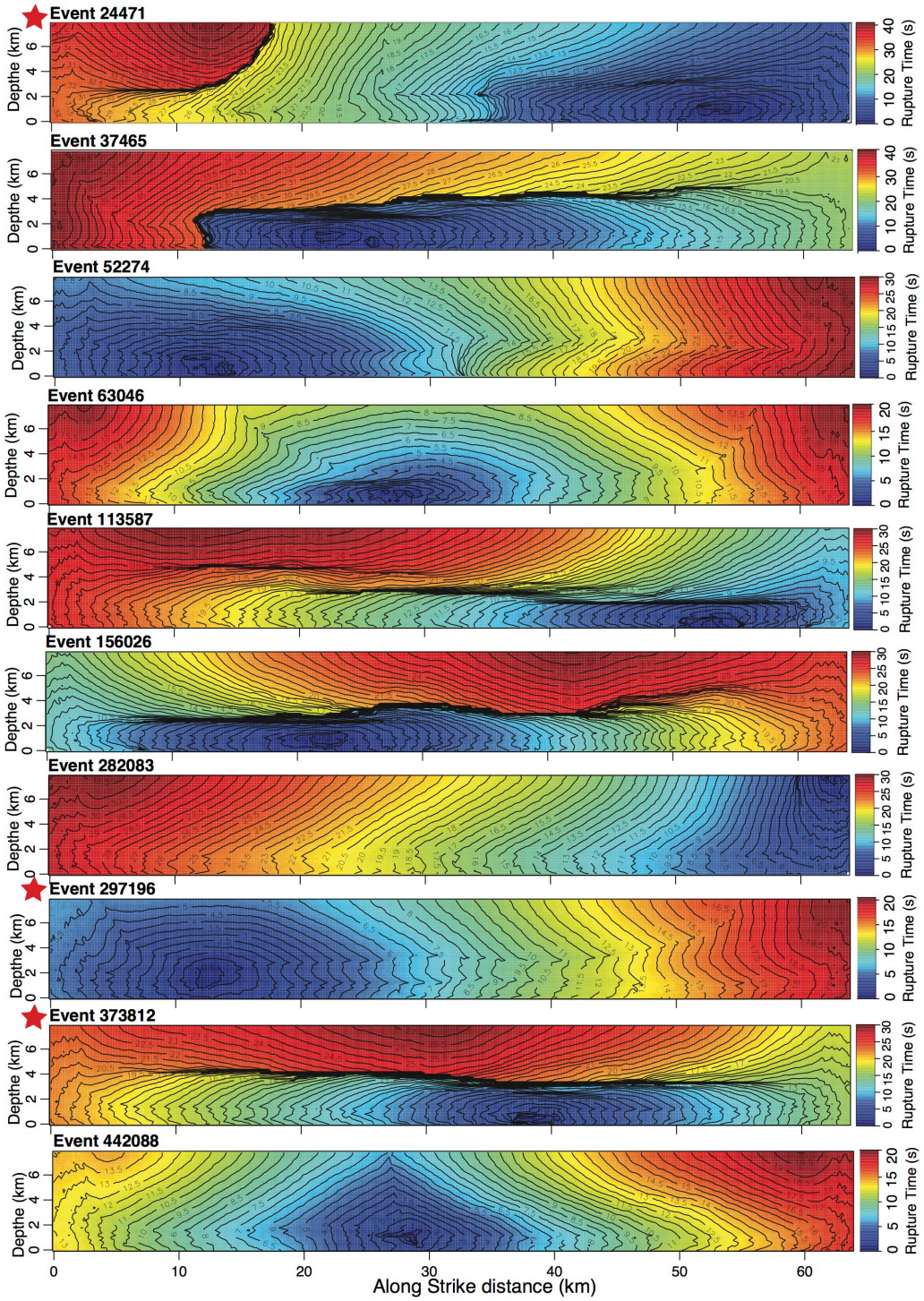

Figure 2.1: Rupture front contours of several events from a 200,000-event RSQSim simulation, with a single, planar, strike-slip fault showing the variation in rupture patterns and velocities of evolved RSQSim events. Black lines show contours of the rupture front as a function of time and color indicates the time at which each section of the fault ruptured. Starred events were selected to test nucleation locations (Figure 2.4, 2.5, &2.6).

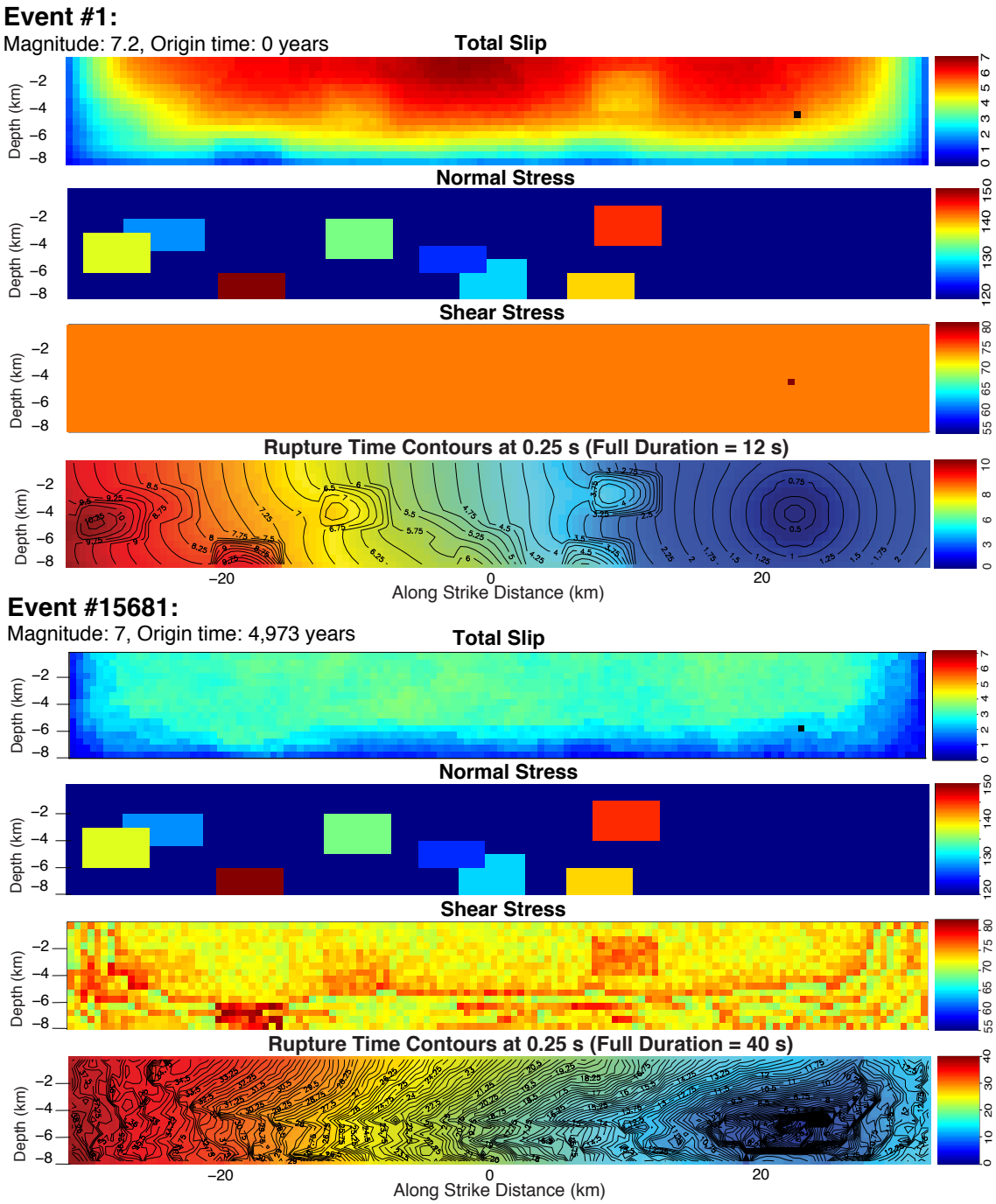

Figure 2.2: Slip, Normal Stress, Shear Stress and Rupture Time Contours of the first event (top) and an evolved event (bottom) of a 200,000-event simulation from RSQSim, with a single, planar, strike-slip fault to illustrate the difference between the first event in RSQSim simulations and the later evolved events. The first event, which was nucleated artificially, was initiated with homogeneous shear stress and heterogeneous normal stress. The evolved event initiated spontaneously with the same non-homogeneous normal stress as the first event (normal stress is constant temporally throughout this simulation), and a heterogeneous shear stress pattern that resulted from fault system interactions over roughly 5,000 years. Note the difference in color scale for the rupture contours.

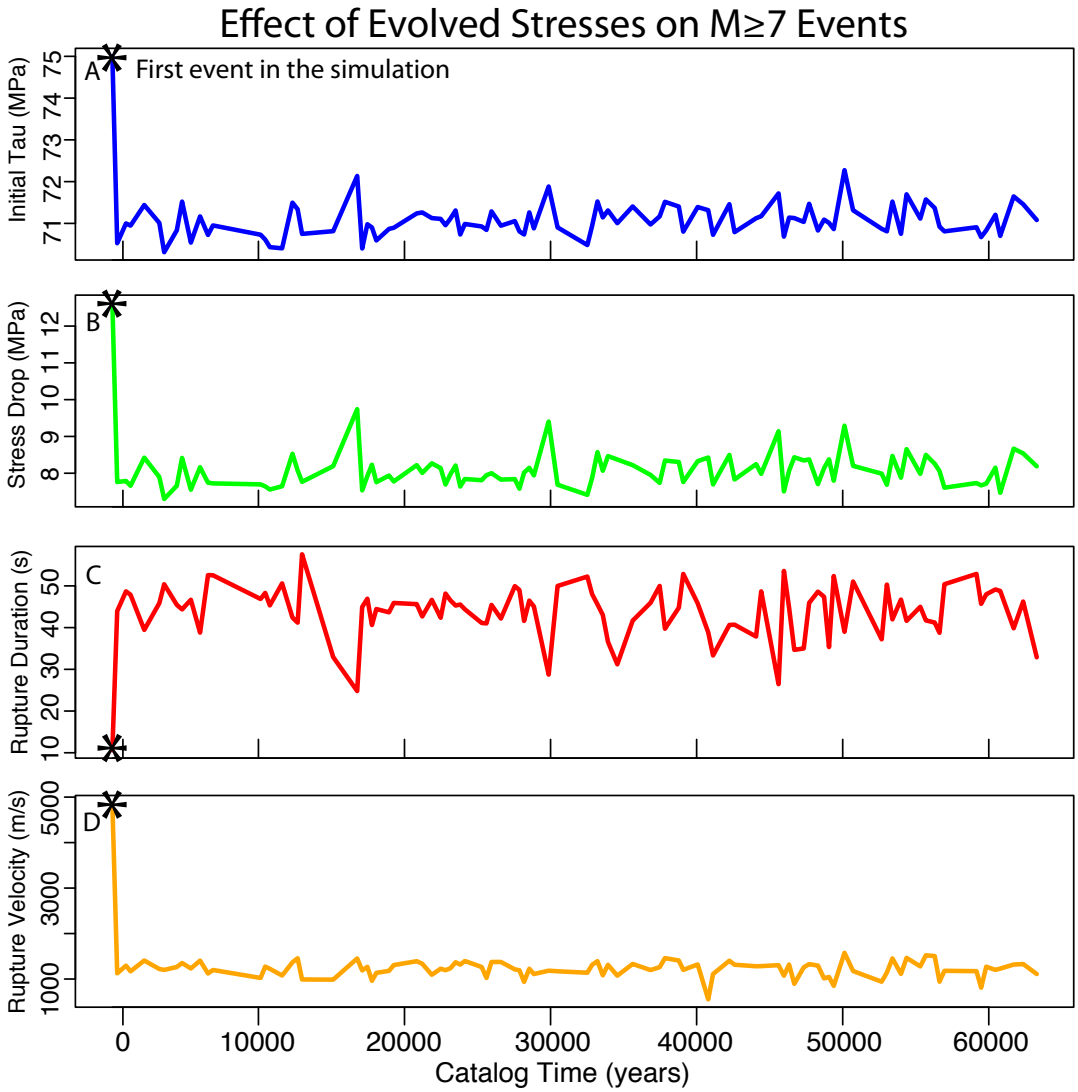

Figure 2.3: Effect of the evolution of shear stress for large (M≥7) events throughout an RSQSim simulation. A) Mean initial shear stress, B) Stress drop, Rupture duration, and D) Rupture velocity. These plots indicate the significant difference in initial shear stress, stress drop, rupture duration, and rupture velocity between the first event in RSQSim simulations and the evolved events.

One important aspect of RSQSim is that the simulations are able to find a minimal stress state for a through-going rupture to propagate. As tectonic stresses increase, this occurs through a process where small events, prior to a through-going event, repeatedly test the ability of a rupture to propagate through the system from failure at different

locations. As such, the simulations are effective in finding a low initial shear stress state and preferred nucleation location to produce a through-going rupture. Additionally, there is no lasting evidence of initial heterogeneities in the rupture patterns due to the redistribution of shear stresses over many earthquake cycles. This characteristic is especially evident in simulations with non-uniform normal stress as shown in Figure 2.2. The strong patches are clearly evident in the slip and rupture time patterns of the first event in the simulation but have almost disappeared from the later event even though the non-uniform normal stress remains unchanged. Lapusta and Liu (2009), found a similar result and showed that a small patch of 20% higher normal stress significantly affected dynamic rupture in the first modeled event, but became "invisible" to dynamic rupture in subsequent events. These results demonstrate the importance of multi-event simulations in understanding the evolution and heterogeneity of fault stresses and consequent effects of added rupture complexity, where rupture propagation slows and sometimes pauses, which will affect ground motions.

2.4.2 Optimal Nucleation Locations

Fault system interactions within RSQSim simulations are also quite efficient in finding the optimal nucleation location, which enables ruptures to propagate at a minimal fault-wide stress condition. RSQSim events nucleate in areas of higher shear stress relative to the steady-state friction (Equation 1.1), but the size of the rupture area, and thus the magnitude of each event, is highly dependent on the stress pattern near the nucleation point. Each RSQSim event nucleates spontaneously with a unique nucleation

location and pattern of heterogeneous initial shear stress that is a function of position on the fault.

To examine the effect of nucleating events in non-optimal locations, the initial stress conditions of three representative events from a long-term RSQSim simulation (those with starred rupture contour plots in Figure 2.1) were used in single-event RSQSim simulations in which nucleation was forced (by reducing the state variable in Equation 1.1 to immediately initiate slip) at different locations along the fault (one simulation for each fault element as the nucleation point – roughly 200,000 simulations in total). These events were selected because their ruptures have a range of patterns representative of different complexities - a simple event (Figure 2.4) with nearly constant rupture speed, a moderately complicated event (Figure 2.5) in which rupture propagation, but not slip, paused for a few seconds near the end of the fault before continuing, and a very complicated event (Figure 2.6) in which the deeper half of the fault ruptured bilaterally before the upper half of the fault ruptured upwards.

For the first set of simulations, nucleation was forced at each fault element in turn, but the initial conditions (the initial shear stress shown in Figures 2.4b, 2.5b and 2.5b) from the original RSQSim event were unchanged. The color contours in Figures 2.4c, 2.5c and 2.6c show the resulting earthquake moment magnitudes resulting from forced nucleation at each element (12,800 simulations each). The vast majority of these simulations failed to propagate across the entire fault and produced events with magnitude 3.4 - 4 (primarily rupturing only the elements where nucleation was forced)

and fewer than 100 nucleation locations (indicated by the deepest red) resulted in magnitude 6.7 - 6.8, end-to-end rupturing events.

For the second set of simulations (Figures 2.4d/e, 2.5d/e, and 2.6d/e) the original pattern of shear stress was preserved, but for each simulation a spatially uniform increment of shear stress $\delta\tau$ (at every fault element) was added until an end-to-end rupture occurred. The new shear stress, across the entire fault is given by:

$$\tau_{New} = \tau_{RSQsim} + \delta\tau \,, \tag{2.2}$$

where $\delta\tau$ is a constant increment of shear stress. This test involved running at least one and often several incremental simulations (until enough shear stress was added that an end-to-end rupture occurred) for nucleation at each of the 12,800 fault elements.

For most forced nucleation locations, (representing nearly 98% of the fault areas), moving the nucleation location away from the point of spontaneous nucleation required increasing the shear stress (often by several MPa) to enable rupture to propagate across the entire fault. These models required added shear stresses *($\delta\tau$)*, of 0.5-3.7 MPa. Figures 2.4d/e, 2.5/e, and 2.6d/e show the amount of added shear stress for each simulation and the resulting stress drop respectively, followed by histograms of each (2.4f, 2.5f, and 2.6f). The additional shear stress lead to stress drops of up to 12 MPa for some of the nucleation locations, which is several times higher than the stress drops of the three original, spontaneous events (~3-4 MPa). Interestingly, the highest stress drops did not always correspond with the highest added shear stresses, but were often related to the pattern of shear stresses and their proximity to the edges of areas with abrupt spatial changes in shear stress. Of the three cases examined, the most complicated RSQSim

event had the fewest locations from which through-going ruptures could nucleate without additional shear stress (Figure 2.6c). On average, the stress drops of these through-going events with non-optimal nucleation locations were 30% to 60% higher than the stress drops of the original RSQSim events, depending on the complexity of the event. The average stress drops of these events vary by a factor of three depending on the nucleation location. Overall, arbitrary nucleation points that are not associated with the pattern of stresses on the fault generate exaggeratedly high stress drops. Stress drop plays a significant role in the magnitude of ground motions (Day, 1982; Oglesby & Day, 2002), therefore variations in stress drop would be expected to result in comparable variations in ground motions.

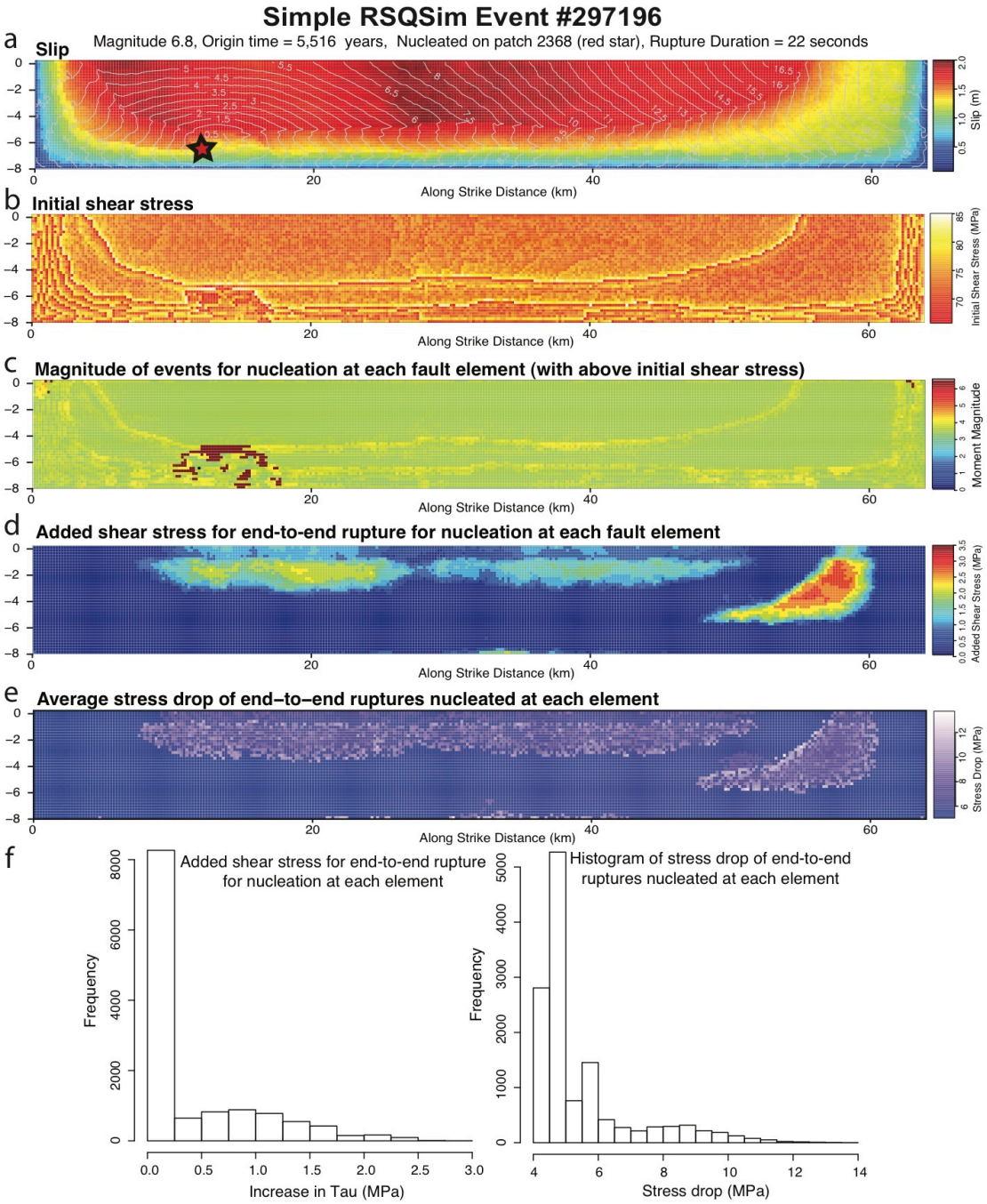

Figure 2.4: Nucleation Test: Simple RSQSim Event. Details from a representative simple RSQSim event from a long-term simulation. Red star indicates the spontaneous nucleation location from RSQSim and color indicates (a) maximum slip and (b) shear stress. Figures c, d, and e represent individual, single-event models for each of the 12,800 fault elements. (c) Magnitude of each event from forced nucleation at that location with original RSQSim shear stress. (d) Amount of added shear stress required for an end-to-end rupture at each location, (e) resulting stress drop for each of those end-to-end events, and (f) histograms of the data from d and e. Nucleating event at non-optimal nucleation locations often requires increasing the shear stress to produce an end-to-end rupture, which results in higher stress drops.

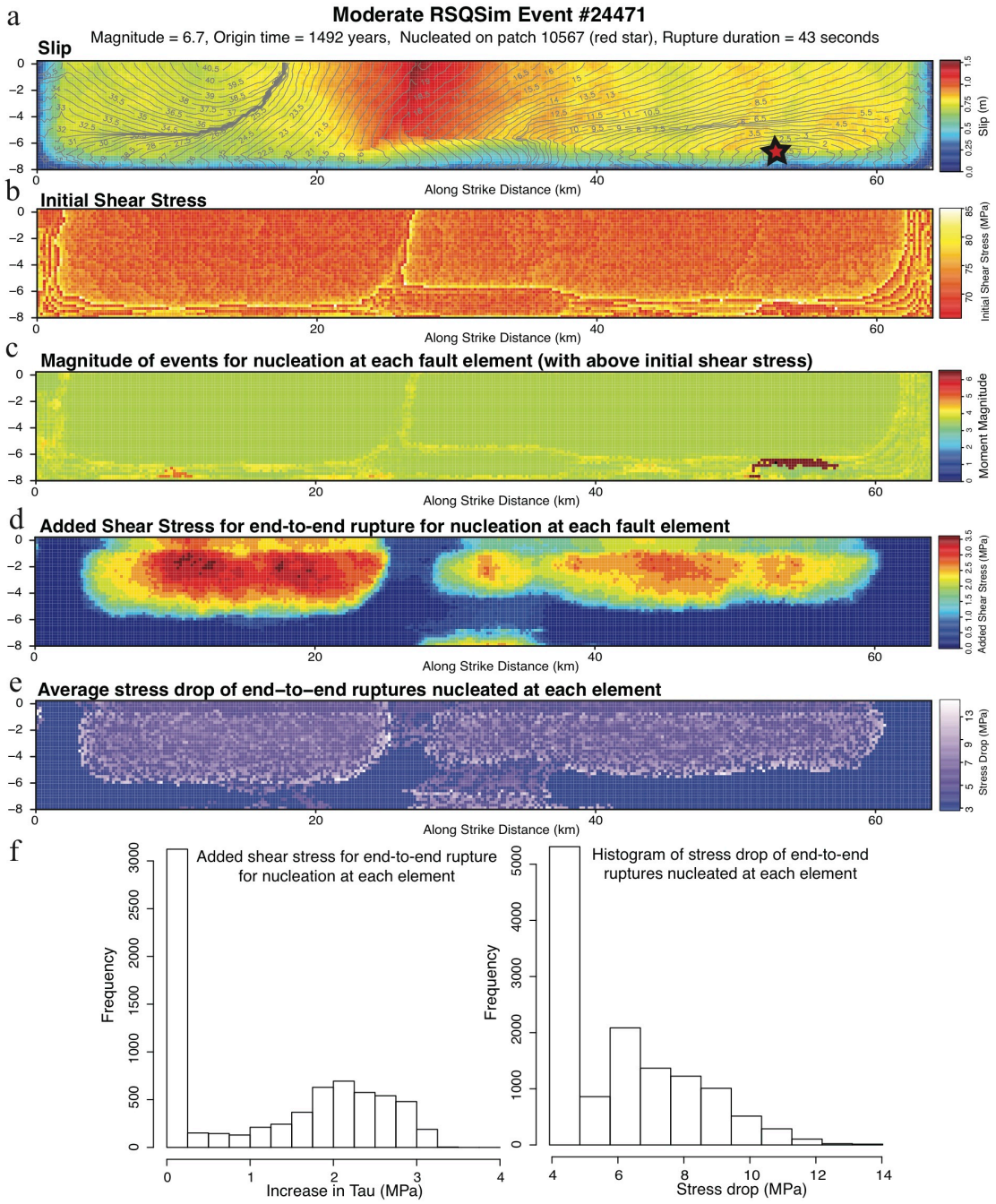

Figure 2.5: Nucleation Test: Moderate RSQSim Event. Details from a representative moderately complicated RSQSim event from a long-term simulation. Red star indicates the spontaneous nucleation location from RSQSim and color indicates (a) maximum slip and (b) shear stress. Figures c, d, and e represent individual, single-event models for each of the 12,800 fault elements. (c) Magnitude of each event from forced nucleation at that location with original RSQSim shear stress. (d) Amount of added shear stress required for an end-to-end rupture at each location, (e) resulting stress drop for each of those end-to-end events, and (f) histograms of the data from d and e. Nucleating event at non-optimal nucleation locations often requires increasing the shear stress to produce an end-to-end rupture, which results in higher stress drops.

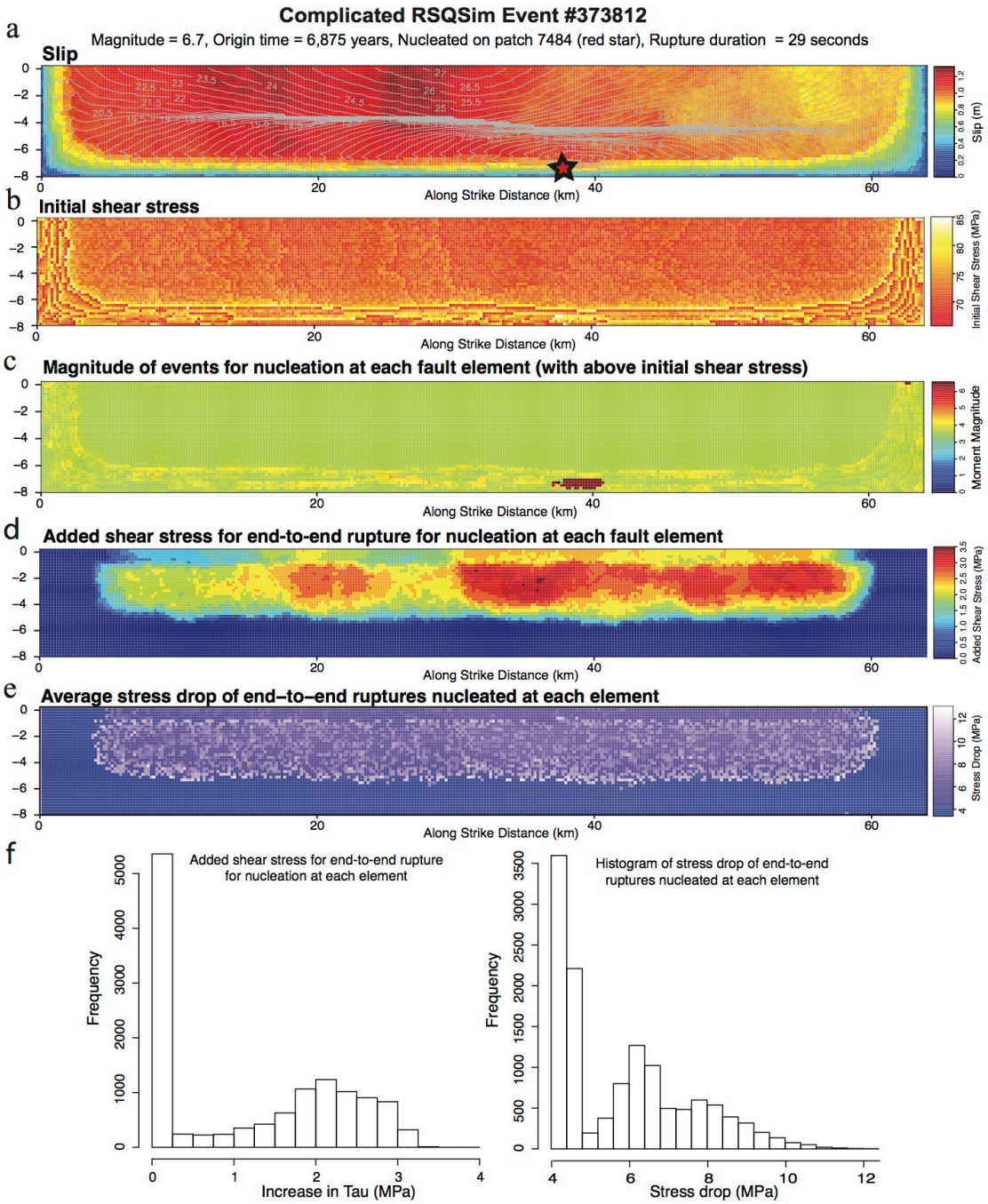

Figure 2.6: Nucleation Test: Complicated RSQSim Event. Details from a representative complicated RSQSim event from a long-term simulation. Red star indicates the spontaneous nucleation location from RSQSim and color indicates (a) maximum slip and (b) shear stress. Figures c, d, and e represent individual, single-event models for each of the 12,800 fault elements. (c) Magnitude of each event from forced nucleation at that location with original RSQSim shear stress. (d) Amount of added shear stress required for an end-to-end rupture at each location, (e) resulting stress drop for each of those end-to-end events, and (f) histograms of the data from d and e. Nucleating event at non-optimal nucleation locations often requires increasing the shear stress to produce an end-to-end rupture, which results in higher stress drops.

This effect is illustrated in Figure 2.7, which gives results from two dynamic rupture simulations. The upper panel in Figure 2.7 gives the peak particle velocity for the rupture simulations with the optimal nucleation location for RSQSim Event #297196 (Figure 2.4a/b), which had an average initial shear stress of 75.4 MPa and a stress drop of 4.1 MPa; and the lower panel gives the results for the non-optimal nucleation location, which required an average initial shear stress of 78.9 MPa which resulted in a stress drop of 6.6MPa. The figures are in map view, with the faults at the centers. Black stars indicate the nucleation points. Ground motions are shown as peak horizontal particle velocity on the surface of the earth for 10km around the fault. With the element size of 200 m near the fault and 400 m away from the fault these models have a minimum frequency of 1.73 Hz – 0.865 Hz, with the assumption that shear waves are well resolved over ten elements (Duan, 2009). The maximum peak particle velocity for the event nucleated at the optimal location was 1.78 m/s, while the event nucleated at the non-optimal location generated a maximum horizontal peak particle velocity of 2 m/s, and velocities in excess of 1.78 m/s over a larger area than the model with the optimal nucleation location. At 5 km from the fault, beyond the areas with the most extreme accelerations, the model with the optimal nucleation location had an average peak particle velocity of 0.60 m/s, while the model with the non-optimal nucleation location had an average peak particle velocity of 0.73 m/s. Additionally, moving the nucleation location greatly changes the directivity pattern of the ground motions. For long, strike-slip faults especially, rupture directivity can cause significantly higher ground motions at the far end of the fault, away from the nucleation location. Moving the nucleation

location changes which areas of the model experience this effect. This becomes particularly important when modeling real world faults, as it could greatly affect the ground motion estimates for communities near major faults. Additionally, models with non-optimal nucleation locations often have higher rupture propagation speeds. Overall, nucleating events at non-optimal locations, and thus at higher initial shear stresses, results in significantly stronger ruptures and higher ground motions over greater areas around the fault, and potentially alters the pattern of these ground motions and which areas of the model which experience the greatest shaking.

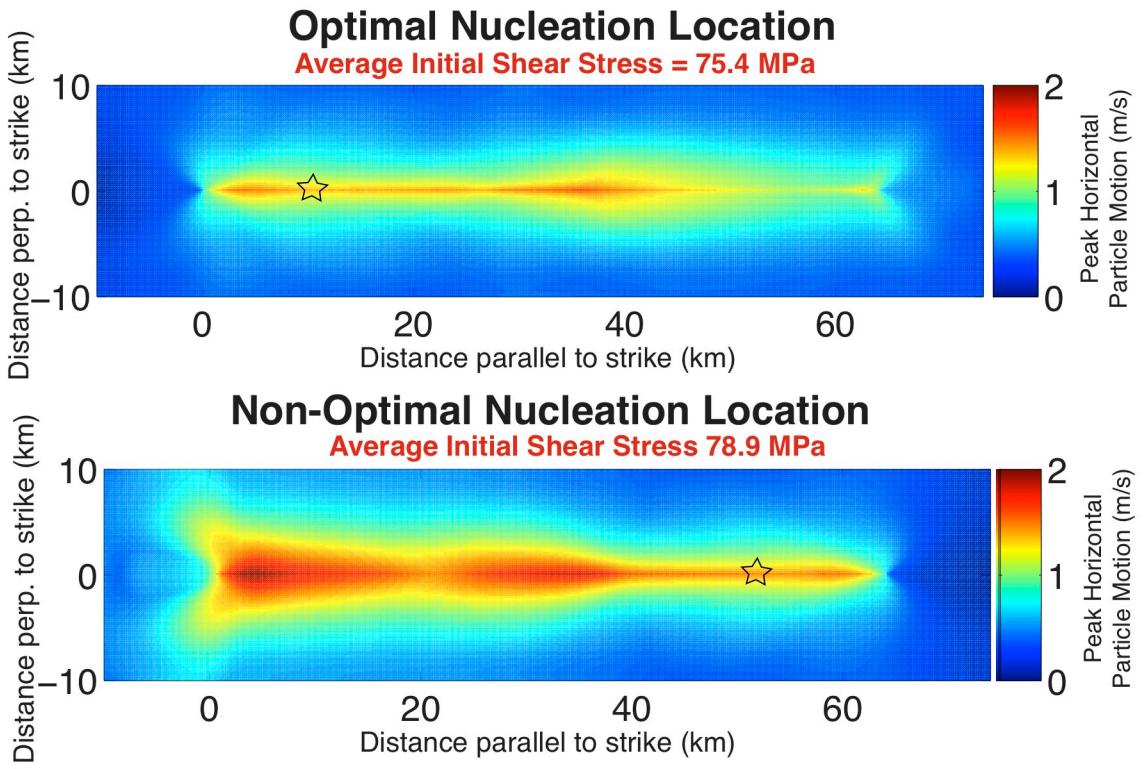

Figure 2.7: Ground motion comparison for optimal and non-optimal nucleation points. Each model is shown in map. Black stars indicate the epicenters. Color indicates the maximum horizontal peak particle velocity for each point on the surface. The shear stresses for the optimal nucleation location (top) are the shear stresses immediately prior to RSQSim Event # 297196 (Figure 2.3) and the shear stresses for the non-optimal nucleation location (bottom) are for the maximum added shear stress (3.5MPa) model from 2.3d. Nucleating events in non-optimal nucleation locations requires increasing the shear stress to produce an end-to-end rupture, which results in higher ground motions.

2.4.3 Effect of Heterogeneous Stresses on Ground Motions

Another series of dynamic fault rupture model tests were performed with FaultMod in order to compare the ground motions of models initiated with heterogeneous stresses to models initiated with homogeneous stresses. These dynamic simulations were performed with a single, planar, strike-slip fault, as before. The results of these simulations indicate that the heterogeneity of initial stresses may significantly affect ground motions. Figure 2.8 shows the peak horizontal particle velocity of four different dynamic models initiated with varying degrees of stress heterogeneity. Measurements of the average peak particle velocity are taken along a parallel profile, 5 km from the fault (PPV5 – listed below each image along with the average stress drop for each model). Model A was initiated with both uniform shear stress (75MPa) and uniform normal stress (123MPa). Model B was initiated with heterogeneous shear stress, in the form of an evolved stress state from a multi-event RSQSim simulation (similar to the shear stress patterns shown in Figures 2.4, 2.5, and 2.6), and with uniform normal stress (123MPa). The ground motions and stress drop that resulted from Model B are weaker than those from Model A. Model C was initiated with uniform shear stress and heterogeneous normal stress in the form of "strong" patches of high normal stress (shown in Figure 2.2). Finally, Model D was initiated with another heterogeneous, evolved shear stress (similar to that of Model B, but from a different multi-event simulation) and the same heterogeneous normal stress as Model C. The ground motions and stress drop of Model C are higher than those of Model D, which has a higher degree of heterogeneity. Models C and D have the same average normal stress and shear stress, suggesting that the

difference in ground motion is due to rupture incoherence, in which the velocity and direction of rupture changes during an event. This coincides with observations that evolved RSQSim events tend to pause or slow down and wrap around barriers on the fault such as the patches of high normal stress shown in Figure 2.2, or in the form of bends, branches, or step-overs in more complex fault systems, such as the UCERF fault models discussed in Chapter 1. Overall, ground motions that result from models with complex, heterogeneous initial stresses from evolved RSQSim events are lower than those from models with simple, homogeneous initial stresses.

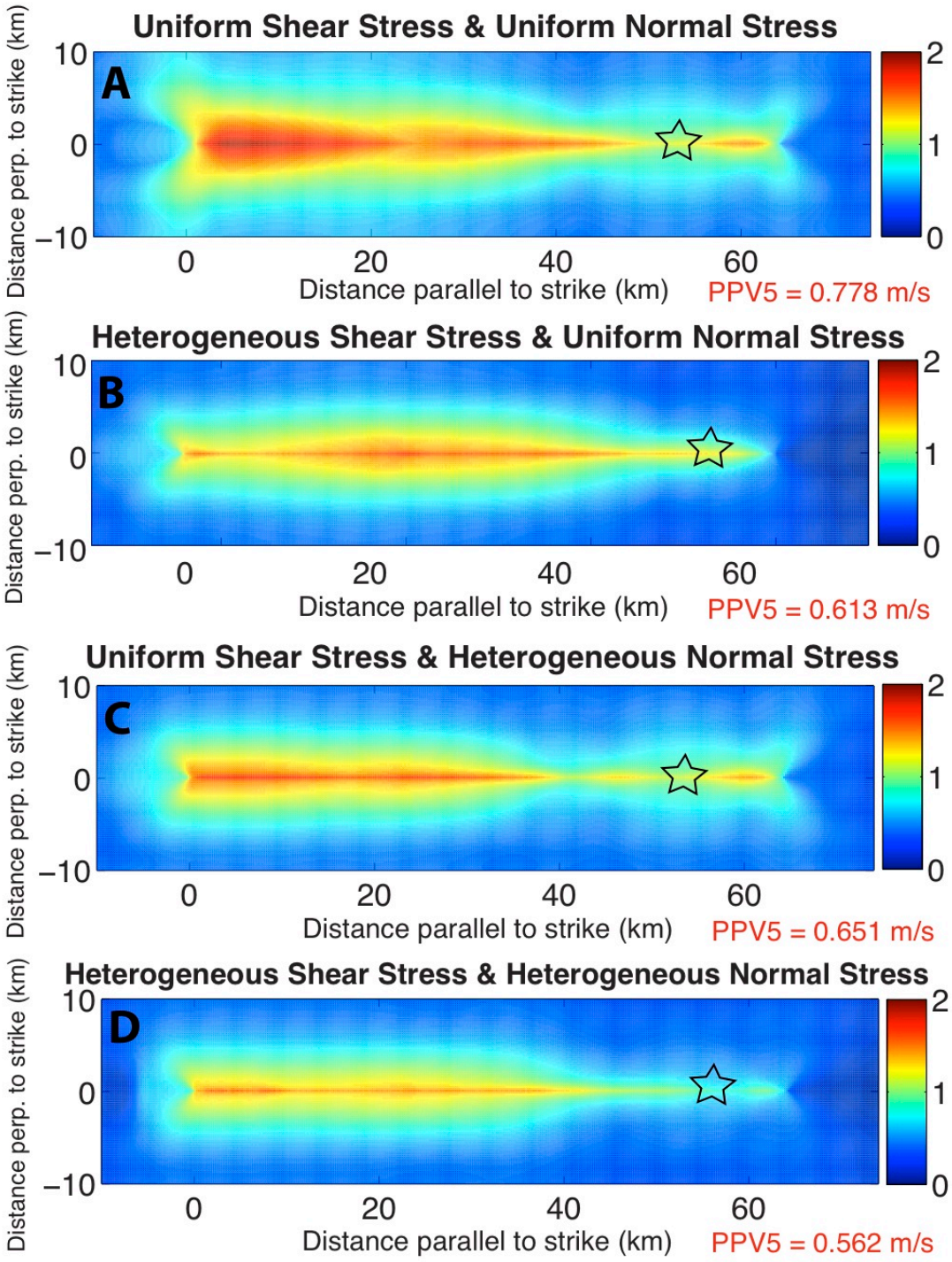

Figure 2.7: Peak horizontal particle motions of four models with different combinations of homogeneous and heterogeneous normal and/or shear stress. Each model is shown in map view with the fault centered horizontally. Black stars indicate the epicenters. Color indicates the peak particle velocity for each point on the surface. The average peak particle velocities along a line at a distance of 5 km from the fault (PPV5) are listed below each figure. Evolved heterogeneous stresses prior to RSQSim events result in weaker ground motions than homogeneous stresses.

2.5 Conclusions

This study explored the characteristics of ruptures in RSQSim simulations with a particular emphasis on how those characteristics might influence ground motions. Fault system interactions within RSQSim simulations produce heterogeneous, evolved stress states that result in a variety of rupture patterns. These fault system interactions are efficient in finding a nucleation location that enables ruptures to propagate at minimal average stress states. Additionally, the evolved stress states that precede large earthquakes in RSQSim result in earthquakes that have weaker ground motions than earthquakes nucleated artificially on faults with homogeneous or prescribed stresses. This is due to the lower stress drops, lower rupture propagation speeds, and less coherent ruptures. On the other hand, forced nucleation at non-optimal locations requires additional shear stress to produce through-going ruptures, which results in higher rupture propagation speeds and ground motions. Although dynamic rupture simulations can be very powerful tools, these results suggest that current methods may not be adequately estimating ground motions. Better initial stress distributions for dynamic models may improve seismic hazard estimates.

Because RSQSim is able to simulate hundreds of thousands to millions of events in large, complicated fault systems, it is feasible to define sets of initial stress conditions for use in generating suites of synthetic seismograms and source models for both kinematic and dynamic rupture simulations. Hence, initial conditions can be generated and used to estimate ground motions for any modeled region or fault system. Specifically, models can be tailored to significant faults in California using long-term simulations with the UCERF

fault models detailed in Chapter 1 and discussed in Chapters 3-5. Synthetic seismograms produced from dynamic rupture simulations are instrumental in the design of earthquake-resilient, critical structures such as bridges and power plants in addition to schools, office buildings, and residences. Better initial stress distributions for dynamic models will improve seismic hazard estimates and allow earthquake scientists, engineers, stakeholders, and policy makers to better evaluate the risk posed by earthquakes that occur in or near populated areas.

2.6 References

Andrews, D. J. (1980). A Stochastic Fault Model 1. Static Case, *Journal of Geophysical Research*, 85, 3867-3877.

Andrews, D. J., and Barall, M. (2011). Specifying Initial Stress for Dynamic Heterogeneous Earthquake Source Models, *Bulletin of the Seismological Society of America*, 101.

Barall M. (2008). FaultMod Finite Element Code, available at **http://scecdata.usc.edu/cvws/download/codedesc/Barall_FaultMod_Desc.pdf**

Candela, T., Renard, F., Bouchon, M., Marsan, D., Schmittbuhl, J., and Voisin, C. (2009). Characterization of fault roughness at various scales: Implications of three-dimensional high-resolution topography measurements. *Pure and Applied Geophysics*, *166*(10-11), 1817-1851.

Cui, Y., Olsen, K. B., Jordan, T. H., Lee, K., Zhou, J., Small, P., ... & Maechling, P. (2010, November). Scalable earthquake simulation on petascale supercomputers. In *Proceedings of the 2010 ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis* (pp. 1-20). IEEE Computer Society.

Day, S. M. (1982). Three-dimensional simulation of spontaneous rupture: The Effect of Nonuniform Prestress. *Bulletin of the Seismological Society of America*, *72*(6), 1881–1902.

Dieterich, J.H., and Kilgore, B.D. (1994). Direct observations of frictional contacts: new insights or state-dependent properties, *Pageoph, 143*, 283-301.

Dieterich, J. H., and Richards-Dinger, K. (2010). Earthquake recurrence in simulated fault systems, *Pure Appl. Geophysics*, **167**,1087–1104, doi:10.1007/s00024-010-0094-0.

Dieterich, J. H., and Smith, D. E. (2010). Nonplanar Faults: Mechanics of Slip and Off-Fault Damage, Pure and Applied Geophysics, 166, 1799-1815.

Duan, B. (2010). Role of initial stress rotations in rupture dynamics and ground motion: A case study with implications for the Wenchuan earthquake. *Journal of Geophysical Research: Solid Earth (1978–2012)*, *115*(B5).

Duan, B., and Oglesby, D. D. (2007). Nonuniform prestress from prior earthquakes and the effect on dynamics of branched fault systems, *Journal of Geophysical Research,* 112.

Harris R.A., Archuleta R.J. (2004). Earthquake rupture dynamics: comparing the numerical simulation methods, EOS, Trans. am. geophys. Un., 85(34), 321.

Harris R.A., et al. (2009). The SCEC/USGS dynamic earthquake rupture code verification exercise, Seism. Res. Lett., 80(1), 119–126, available at **http://scecdata.usc.edu/cvws/**.

Lapusta, N., and Liu, Y. (2009), Three-dimensional boundary integral modeling of spontaneous earthquake sequences and aseismic slip, J. Geophysics. Res., 114, B09303, doi:10.1029/2008JB005934.

Linker, M.F., and Dieterich, J.H. (1992). Effects of variable normal stress on rock friction: observations and constitutive equations, *Journal of Geophysical Research, 97*, 4923-4940.

Lozos, J. C., Olsen, K. B., Brune, J. N., Takedatsu, R., Brune, R. J., & Oglesby, D. D. (2015). Broadband ground motions from dynamic models of rupture on the northern San Jacinto fault, and comparison with precariously balanced rocks. *Bulletin of the Seismological Society of America*, *105*(4), 1947-1960.

Mai, P. M., Spudich, P., and Boatwright, J. (2005). Hypocenter Locations in Finite-Source Rupture Models, *Bulletin of the Seismological Society of America*, 95, 965-980.

Mai, P. M., Imperatori, W., & Olsen, K. B. (2010). Hybrid broadband ground-motion simulations: Combining long-period deterministic synthetics with high-frequency multiple S-to-S backscattering. *Bulletin of the Seismological Society of America*, *100*(5A), 2124-2142.

Mena, B., Mai, P. M., Olsen, K. B., Purvance, M. D., & Brune, J. N. (2010). Hybrid broadband ground-motion simulation using scattering Green's functions: Application to large-magnitude events. *Bulletin of the Seismological Society of America*, *100*(5A), 2143-2162.

Oglesby, D. D., Mai, P. M., Atakan, K., and Pucci, S. (2008). Dynamic models of earthquakes on the North Anatolian fault zone under the Sea of Marmara: Effect of hypocenter location, Geophysics. Res. Lett., 35, L18302, doi:10.1029/2008GL035037.

Oglesby, D. D., and Day, S. M. (2002). Stochastic fault stress: implications for fault dynamics and ground motion, *Bulletin of the Seismological Society of America*, 92, 3006-3021.

Okubo, P. G., and Dieterich, J. H. (1984), Effects of physical fault properties on frictional instabilities produced on simulated faults, J. Geophysics. Res., 89(B7), 5817–5827, doi:10.1029/JB089iB07p05817.

Power, W. L., Tullis, T. E., Brown, S. R., Boitnott, G. N. , and Scholz, C. H. (1987). Roughness of natural fault surfaces, *Geophysics. Res. Lett.,* 14, 29 – 32, doi:10.1029/GL014i001p00029.

Sagy, A., and Brodsky, E. E. (2009). Geometrical and rheological asperities in an exposed fault zone, *Journal of Geophysical Research,* 114, 2009.

Chapter 3

How do Models of Paleoseismic Rupture Detectability Affect Estimates of Earthquake Probabilities?

3.1 Abstract

One application of RSQSim for gaining a better understanding of earthquake hazard is the estimation of earthquake probabilities. Estimation of earthquake probabilities relies heavily on estimates of mean recurrence intervals from paleoseismicity. However, paleoseismic records are often considered incomplete due to the difficulty of detecting ruptures in trenches. Alternatively, misinterpretation of paleoseismic features could result in over-counting, which may be a more significant problem for computing long-term mean recurrence intervals and earthquake probabilities. To test the effects of event detectability in paleoseismic studies on the probabilities of large earthquakes in California, several simulated catalogs were tuned to match the reported recurrence intervals at the paleoseismic sites in the UCERF3 report assuming under-, complete-, and over-detection. The time-dependent probabilities of large earthquakes in California were computed for all of these simulated catalogs at each of the paleoseismic sites. Comparisons between the different catalogs suggest that the mean recurrence intervals from paleoseismic trenches may be too short, possibly due to over-counting of events, resulting in over-estimation of probabilities.

3.2 Introduction

3.2.1 Earthquake Probability Calculations

One application of earthquake simulators for gaining a better understanding earthquake hazard is the estimation of earthquake probabilities. There are two types of earthquake probability calculations: time-independent probabilities, which are based on estimates of long-term earthquake rates; and time-dependent probabilities, which take into account the time since the last rupture, and thus how much time stress has been accumulating (Cornell, 1968). The latter method requires estimates of the mean recurrence interval (MRI), the coefficient of variation of the recurrence times (COV), and the elapsed time (Tmre) since the most recent earthquake (Field et al. 2013). The COV is calculated from the standard deviation of the recurrence times divided by the mean recurrence time. The main datasets for MRI and COV come from paleoseismic investigations, which involve trenching across faults to observe the effects of past ruptures, and dating the material to get an estimate of the timing of those ruptures. Probability estimates are used to generate earthquake rupture forecasts. Figure 3.1 shows the four main components of an earthquake rupture forecast as used in the Unified California Earthquake Rupture Forecasts Versions Two (UCERF2) and Three (UCERF3) from Field et al. (2008, 2013). The first step consists of assembling the fault model for California, which comprises the known major, active faults thought to be capable of generating damaging earthquakes. The second step entails assigning long-term slip rates to each of the faults in the model. The third step involves calculating the long-term rates of all potentially damaging earthquakes using fault models, which specify the size,

locations, and geometry of the faults, and deformation models, which specify the slip rates of the faults in the model. For the UCERF2 and UCERF3 studies, this step involved a "Grand Inversion" that is constrained by the fault and deformation models together with various empirical scaling relations. Finally, the probability that each possible earthquake will occur during some specified time span is calculated from the long-term rate models.

In the present study, the fault and deformation models from the UCERF3 report (Field et al., 2013) were used as inputs to RSQSim. Then the models were tuned to the UCERF3 mean recurrence interval to directly generate RSQSim earthquake-rate models, which can be used to calculate time-dependent earthquake probabilities.

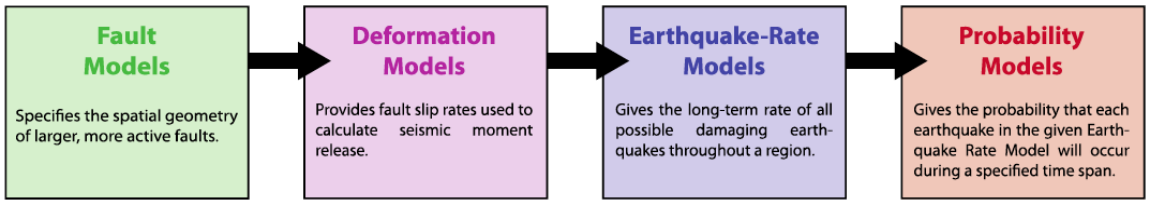

Figure 3.1: UCERF3 Earthquake Rupture Forecast Components (Field et al., 2013, Figure 2, p. 7). Fault and Deformation models are used as inputs to RSQSim, and long-term earthquake rate models are produced from RSQSim simulation.

The standard method for determining conditional earthquake probabilities is to calculate the probability for an event (in some specific magnitude or moment range) at a specific location conditional on the elapsed time (Tmre) since the last event in that magnitude range (McCalpin, 2009). This is the probability that the next rupture at that site will occur in some given amount of time in the future (typically Δt = 30 years) conditional on the elapsed time since the prior event at that site. To compute a probability, this method requires specification of the probability density for recurrence of the defined events as a function of the elapsed time since the last event. This is typically

done with a probability density function (PDF), such as log normal or Brownian passage time, with a specific mean recurrence interval and a measure of the aperiodicity such as the COV. Rather than using a functional representation for the recurrence-time probability density distributions, this study uses empirical PDFs derived from RSQSim catalogs (Dieterich and Richards-Dinger, 2010). This probability can be illustrated graphically from an empirical PDF of the recurrence intervals of events at a paleoseismic site (McCalpin, 2009). An example is shown in Figure 3.2.

The conditional probability can be calculated from the recurrence interval PDF by the following:

$$P_c = A/(A+B), \tag{3.1}$$

where A is the probability of an event in the interval, Tmre to Tmre + Δt, and B is the probability of an event in the open interval Tmre + Δt to T_∞.

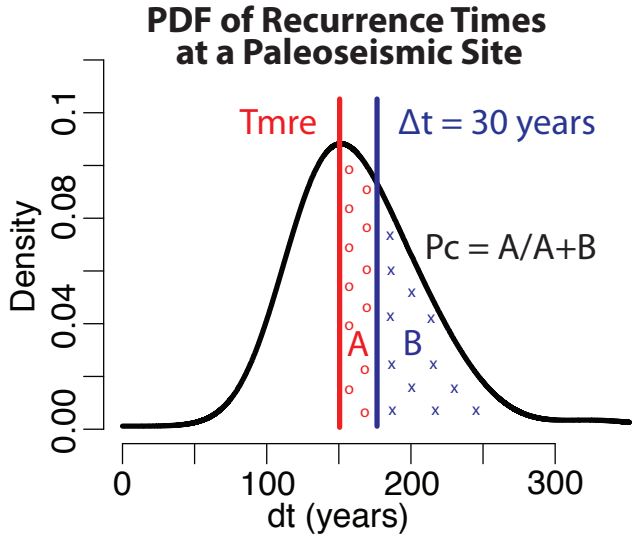

Figure 3.2: Cartoon of a probability density function of the recurrence intervals for events at a paleoseismic site used to calculate long-term conditional probabilities. The red line indicates the time since the most recent event at that site (Tmre) and the blue line indicates the amount of time in the future (30 years in this example) for which the conditional probability (P_c) is to be calculated. A is equal to the area under the curve from Tmre to Tmre + Δt, and B is equal to the entire remaining area under the curve after Tmre.

The first step requires tuning the RSQSim catalogs to the mean recurrence intervals of the paleoseismic sites in California. The UCERF3 report gives the location, long-term MRI and the uncertainties in those calculations, the Tmre, and the mean long-term rate for the thirty-two paleoseismic sites across California. The MRI is estimated at each site using paleoseismic records, including the number of events that ruptured each site and the estimated dates for each event. Once the RSQSim catalogs were tuned to this dataset (the method for which is described below in section 3.3), 30-year conditional probabilities were calculated based on the empirical PDFs for each UCERF3 paleoseismic site.

3.2.2 Paleoseismic Recurrence Intervals in California and the Modern Earthquake Hiatus

Paleoseismic records are often considered likely to be incomplete due to the difficulty of detecting ruptures in trenches. Events become increasingly difficult to detect at smaller amounts of slip, large earthquakes may be missed if the rupture bypassed the portion of the fault that was trenched, or the sedimentation rate could have been low at the time of an event, and therefore not conducive to constraining ruptures. It is also difficult to differentiate between events that occur very close in time. One might assume that these difficulties would result in missing or fewer events. Alternatively, misinterpretation of features in paleoseismic trenches could result in over-counting. This may be a more significant problem for computing long-term mean recurrence intervals

and earthquake probabilities. For example, using UCERF3 data, David Jackson (SCEC 2015, Abstract #081) estimated that the probability of a 100-year or longer earthquake hiatus, such as the one that may be occurring now, is 1% or less. This very low probability may indicate a problem with the reported paleoseismic site MRIs. At the time of publication, eleven of the thirty-two UCERF3 paleoseismic sites are 'overdue', i.e. the time since the most recent event at those sites was past the mean recurrence interval for those sites, and two others (Burro Flats and Plunge Creek) were within a few years (~5) of passing their mean recurrence intervals. If the reported MRIs are correct, the probability of an impending large Southern San Andreas Fault earthquake would be quite high. An alternative possibility is that the earthquake rates in California are not, at least not currently, as high as the paleoseismic data suggest.

There are several possible explanations for this apparent hiatus in modern earthquakes. The first, but least preferred explanation is simply coincidence that there has not been an earthquake at any of those sites. The second explanation is that the modern large-event rates are lower than that implied by paleo-event rates, possibly due to stress shadowing from the last major Northern and Southern San Andreas events, the M7.8 San Francisco Earthquake in 1906 and the M7.9 Fort Tejon Earthquake in 1957 (Simpson and Reasenberg, 1994; Harris and Simpson, 1998). The third possible explanation is that interpretation of event rates from some of the paleoseismic sites are too high, perhaps due to over-counting of events in paleoseismic trenches. This could arise from misinterpretation of geologic features in paleoseismic trenches, such as triggered surface slip from a distant earthquake, landslides, or multiple rupture splays being interpreted as

individual events, and could inflate paleo-event rates compared to modern-event rates that are based on instrumentation and/or direct observation and thus should not include such errors. In order to test these hypotheses we investigated the effect of paleoseismic event detectability on the recurrence intervals of simulated events using several tuned RSQSim catalogs. Conditional probabilities for large events in these catalogs are then compared to determine the effect that different levels of paleoseismic detectability might have on estimates of earthquake probabilities.

3.2.3 California Fault Model and Paleoseismic Sites

The RSQSim-UCERF3 fault model includes all of the UCERF3 faults, discretized with roughly 260,000, 1 km^2, triangular fault elements. The shear stressing rates for this model were set using the backslip method (described in Ch. 1, Section 1.3) to match the long-term geologic slip rates from the UCERF3 report. The UCERF3 fault model, colored by the long-term slip rate, is shown in Figure 3.3. This new California fault model was updated to include all of the faults added to the UCERF2 model in the UCERF3 report and any changes in long-term slip rates. Additionally, the fault elements are significantly smaller at 1 km^2 than the UCERF2 fault model, which was discretized with roughly 12,500, 9 km^2 rectangular elements. This not only reduced the minimum magnitude to roughly magnitude 4 from magnitude 4.9, but the triangular elements fit the curved and dipping faults better than the rectangular elements.

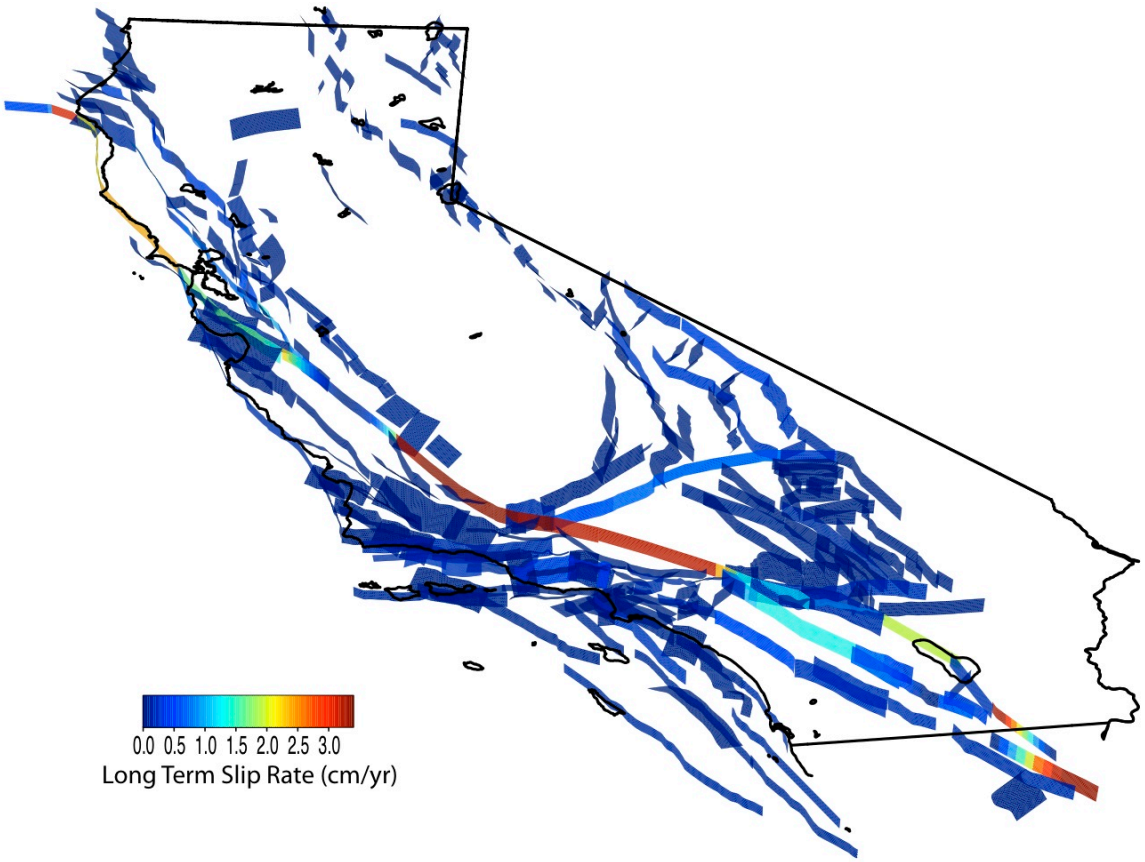

Figure 3.3: Entire UCERF3 Fault Model used to simulate RSQSim catalogs colored by the long-term slip rate. Roughly 260,000, 1-km^2, triangular elements.

The surface traces of the faults are shown in Figure 3.4 with the locations of all of the UCERF3 paleoseismic sites. Each site is assigned to the single fault element that is nearest to the coordinates of the paleoseismic site (UCERF3 Appendix H, Table H3). Although all of the sites are shown in Figure 3.4, no fault element and no recurrence data were assigned to site #17, which is indicated as an offshore core locality interpreted as containing turbidite evidence for multiple past earthquakes. Furthermore, because this study focuses on event detectability at the specific UCERF3 paleoseismic sites, for tuning

purposes, only events that rupture at least one of those thirty-one fault elements are considered. An example of a simulated event used for tuning is shown in Figure 3.5. In this example, a Magnitude 7.4 occurred on the Mojave Section of the Southern San Andreas Fault, rupturing the Pallett Creek, Wrightwood, and Pitman Canyon paleoseismic sites (indicated by black hexagons). The fault elements that ruptured in this event are colored by the total slip at each element. It is not uncommon for several concurrent sites to rupture in the same event, particularly for this long section of the Southern San Andreas Fault. Additionally, Figure 3.5 shows a close-up of several Southern California faults in which the geometric complexity of the fault system and small size of individual fault elements is evident.

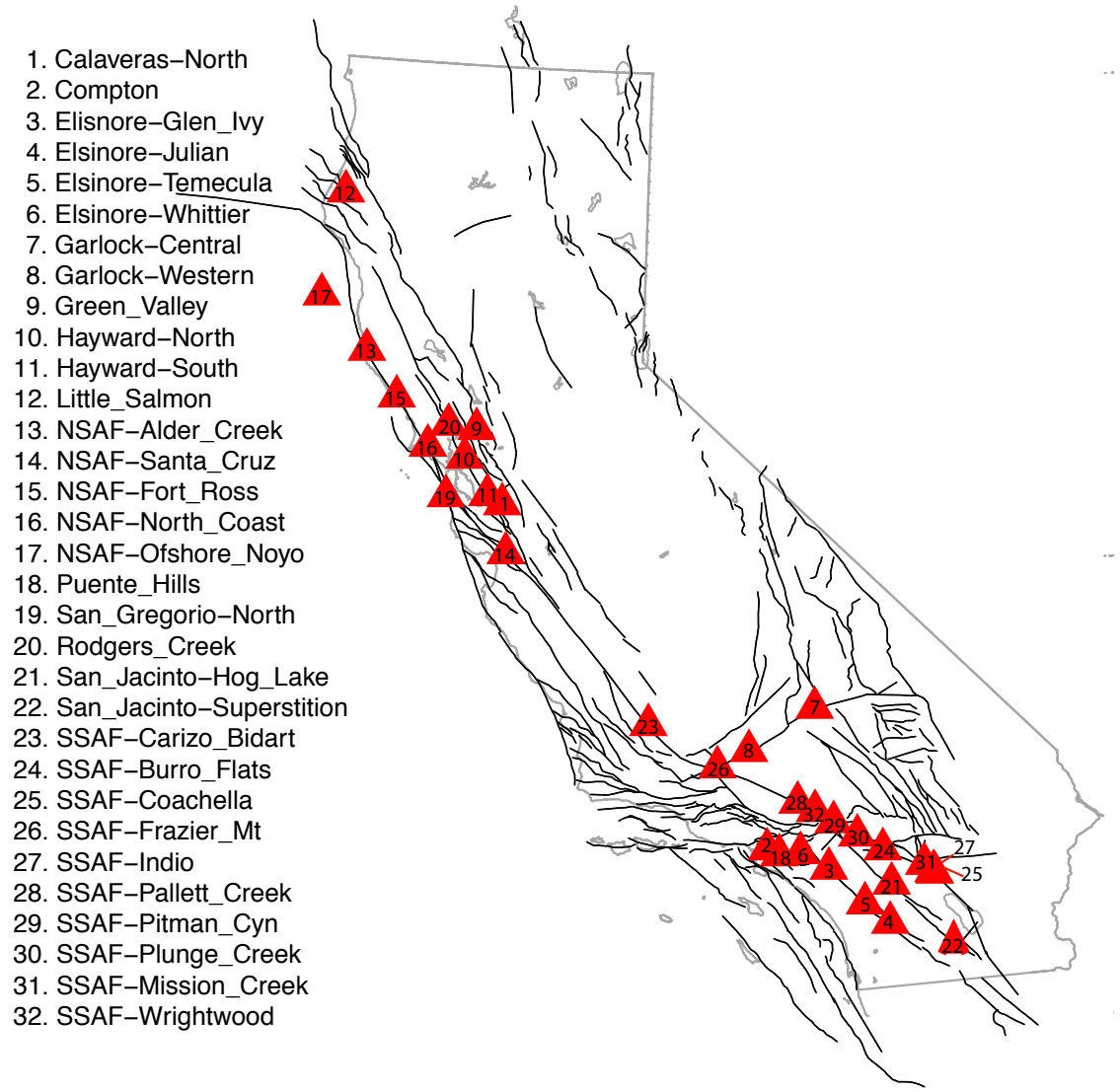

Figure 3.4: Locations of all 32 UCERF3 Paleoseismic Sites (red triangles) mapped onto the surface traces of the UCERF3 Faults. These sites were used to sort RSQSim events and tune the UCERF3 catalogs to the reported mean recurrence interval for each site.

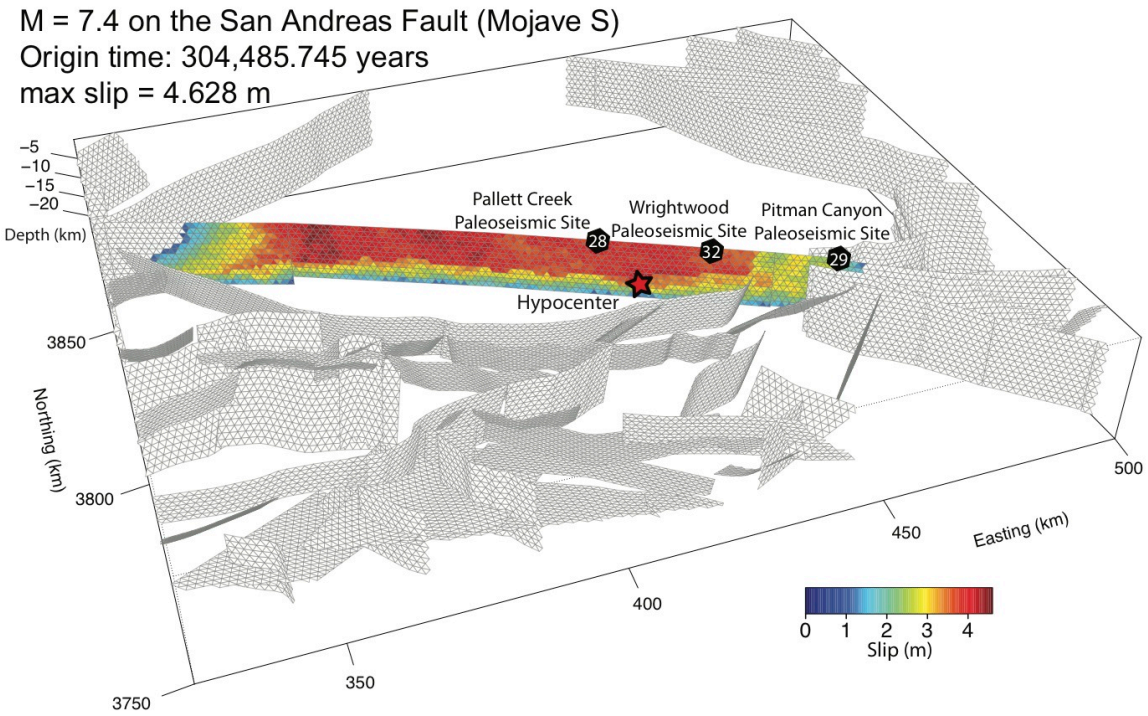

Figure 3.5: Example of an event that ruptured a paleoseismic site (in this case three sites). This event was a M7.4 on the Mojave Section of the Southern San Andreas. Fault elements are colored by the total amount of slip for that event. A red star indicates the hypocenter and the paleoseismic sites that the event ruptured through are indicated by black hexagons with the site number in white.

3.3 Tuning RSQSim California Catalogs

Four different catalogs were tuned to different models of paleoseismic detectability. These catalogs are 1. Under-Detection 'Paleoseismic' catalog, 2. Under-Detection 'Instrumental' catalog, 3. Complete-Detection catalog, and 4. Over-Detection catalog. Each catalog was thinned of events according to the detection model used, prior to tuning, but the specific thinning process was slightly different for each catalog. Descriptions of these catalogs and their differences are detailed in Section 3.3.2 and the general tuning process is described below.

3.3.1 Catalog Thinning and Tuning Steps

During the tuning process, the catalogs were thinned, i.e. events were removed from the catalog if they were not 'detected', based on the specified detectability of each catalog. A flow chart of the tuning steps is shown in Figure 3.6. The first step of the tuning process was to run a moderate length RSQSim simulation (roughly 100,000 years) using the UCERF3 faults, geologic slip rates, and uniform normal stress (100 MPa) on all elements. Second, events in the catalog were sorted by paleoseismic site and thinned based on several requirements. All events that ruptured ≤10 fault elements were thinned from the catalogs, as these RSQSim events of M≤4 tend to have excessive slip for their rupture area. This is in order to avoid over-counting events because in nature, slip in those events is highly unlikely to rupture to the surface. Next, pairs or clusters of events that occurred too close in time to be distinguished as separate events in a paleoseismic trench, i.e. 15 years or less (UCERF3 Appendix I), were combined into single events and assigned the event time of the first event in the series. Then, events were thinned-out based on their probability of detection as a function of slip (Figure 3.7). The UCERF3 probability of detection model assigns an increasing probability of detection for events with greater slip because larger slip events are more likely to be detected in a trench. Note that not all of the thinning requirements were applied to each catalog. The details and explanations for the specific thinning process for each catalog are outlined in Section 3.3.2. Third, once the catalog was thinned to only include events that would have been 'detected' based on the specified model of detectability, the mean recurrence intervals of

the thinned catalogs were compared to the UCERF3 paleoseismic recurrence intervals (UCERF3 Appendix H, Table H3). Fourth, if the simulated mean was not within the 68% confidence intervals of the paleoseismic mean (however, most were adjusted to be well within this interval, usually within a few years of the paleoseismic mean recurrence interval), the normal stress in the RSQSim model was adjusted accordingly in order to shift the simulated recurrence intervals for events at each paleoseismic site towards the mean recurrence interval goal for that catalog. This was done in terms of a Stress Adjustment Factor (F) that is a function of the Mean Recurrence Interval Ratio (R) and a Spatial Weighting Factor (S). The ratio R tells us how much the simulated MRI needs to change to match the paleoseismic MRI and is determined by

$$R = \frac{\mathrm{MRI_{paleo}}}{\mathrm{MRI_{sim}}}, \tag{3.2}$$

where $\mathrm{MRI_{paleo}}$ is the reported paleoseismic mean recurrence interval and the $\mathrm{MRI_{sim}}$ is the RSQSim mean recurrence interval for each paleoseismic site. The spatial weighting factor S varies for each fault element and assigns a greater weight to patches that slipped more in the events that ruptured the paleoseismic sites. It is determined by

$$S_i = \frac{Slip_i}{Slip_{max}}, \tag{3.3}$$

where S_i is the sum of the slip on the ith fault element ($slip_i$) in all events that ruptured that element, normalized by the maximum slip ($slip_{max}$) in all events that ruptured that element. The stress adjustment factor is then determined by

$$F_i = 1 + \frac{S_i \times (R-1)}{S_{total}}, \tag{3.4}$$

where F_i is the stress adjustment factor on each fault element *i,* S_i is the spatial weighting factor for fault element *i,* R is the mean recurrence interval ratio for that paleoseismic site, and S_{total} is the sum of the slip for all events that ruptured that fault element. We then define an adjusted normal stress

$$\sigma_i^{adj} = F_i \times \sigma_0, \tag{3.5}$$

where σ_i^{adj} is the adjusted normal stress on each fault element in the model for the next simulation, σ_0 is the normal stress on each fault element in the model from the previous simulation, and F_i is the Stress Adjustment Factor.

Adjusting the normal stress in the model is a very effective way to shift the MRIs of events in the simulations. Increasing the normal stress on a fault effectively increases the frictional strength of the fault, which results in less frequent ruptures, thus increasing the recurrence interval for events on that fault. The reverse is also true. The area of the fault over which these normal stress adjustments were made was determined by the extent of the ruptures of events that were detected at each paleoseismic site. The spatial weighting factor (S) was used to taper the normal stress adjustments and avoid sharp boundaries in fault strength. The spatial weighting factor was set to 1 at each paleoseismic site. Some events ruptured more than one paleoseismic site, resulting in elements that had normal stress adjustments from more than one paleoseismic site. If this was the case, a weighted average was used to give more weight to the adjustments from the nearest paleoseismic site to that element. Next, a new RSQSim simulation was run with the adjusted normal stresses and the tuning process was repeated. Finally, once the RSQSim mean recurrence intervals matched those of the current goal (after roughly 3 to

4 tuning iterations) a long (>500,000 years) simulation was run with the tuned parameters. Figure 3.8 illustrates the change in mean recurrence interval for the paleoseismic sites in the Under-Detection 'Paleoseismic' catalog over four tuning iterations. Due to variation in recurrence times in the shorter catalogs used for tuning, later tuning iterations would sometimes slightly over- then under-shoot the exact paleoseismic MRI, but all final MRIs were within the 68% confidence bounds of the UCERF3 paleoseismic MRIs (Field, et al., 2014).

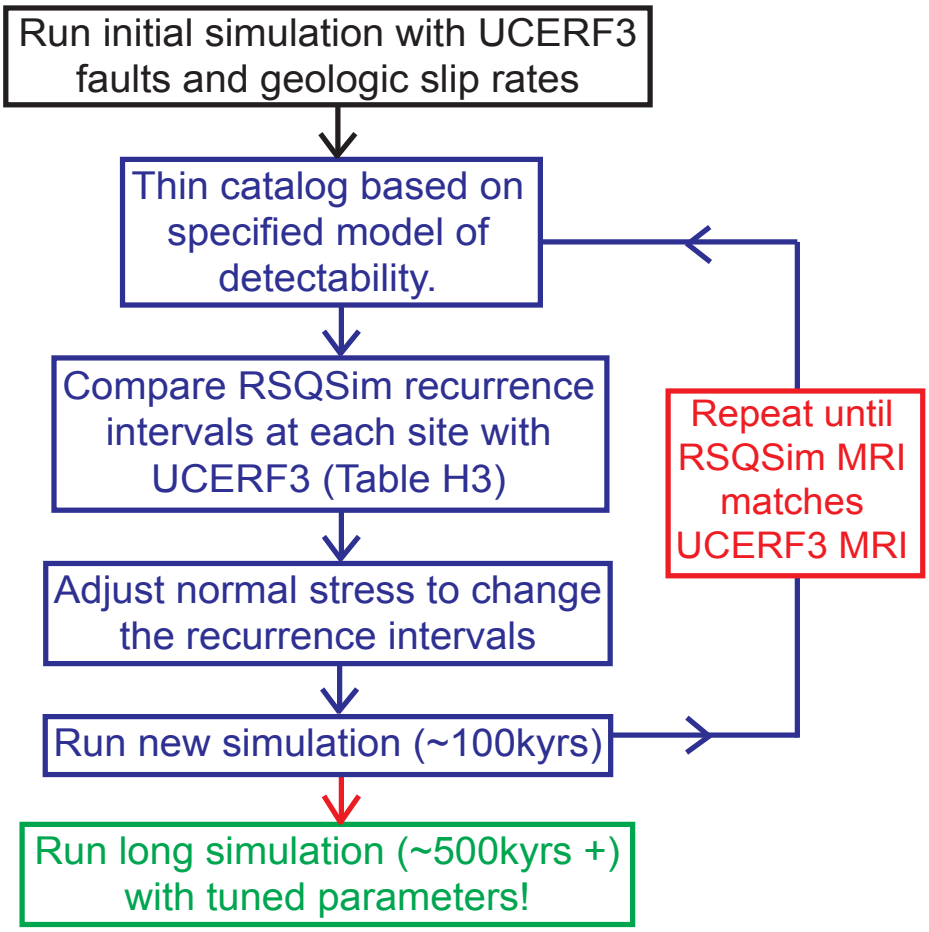

Figure 3.6: Flow chart showing the steps in the tuning process for the simulated California catalogs. The tuning process involves a short initial simulation, several tuning iterations (for which the normal stress is adjusted to change the recurrence intervals), and a final long simulation with the tuned parameters.

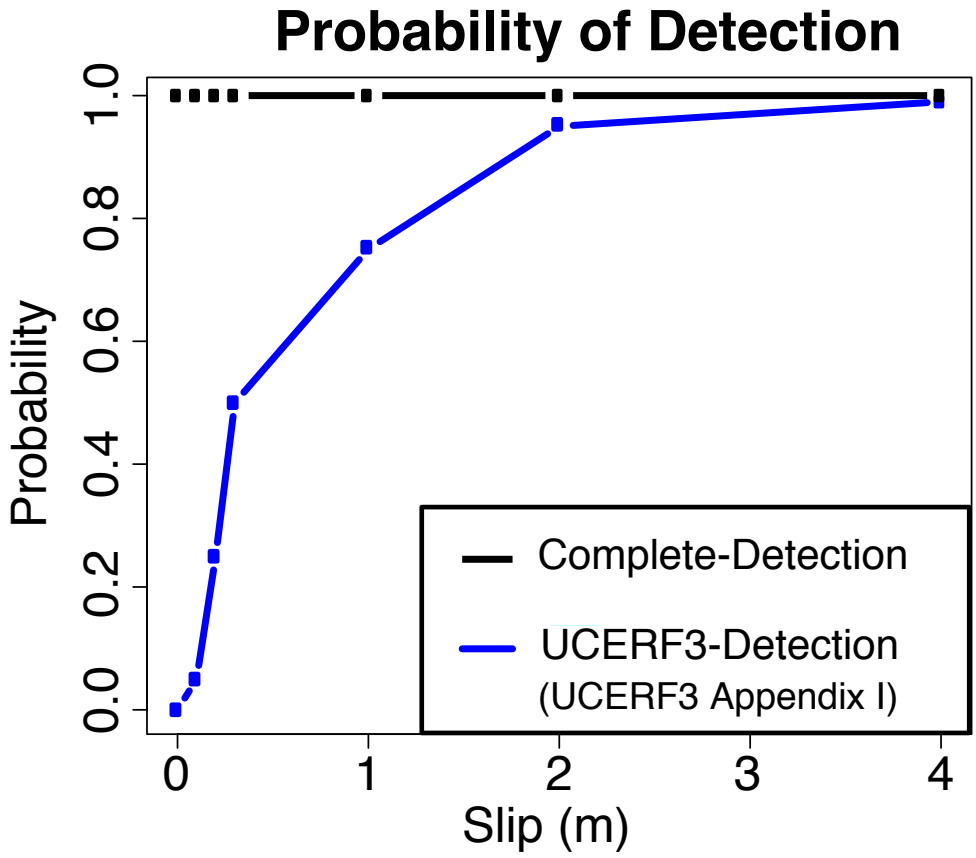

Figure 3.7: The two different probability of detection functions used in the thinning steps of both Tuning and Processing of the UCERF3 catalogs. Both functions specify a probability of detection for events at the paleoseismic sites based on the amount of slip at that site. The UCERF3 Detection Function (Appendix I), shown in blue, assumes that events with greater amounts of slip are more likely to be detected, while the Complete Detection Function (shown in black) assumes 100% detection for all amounts of slip.

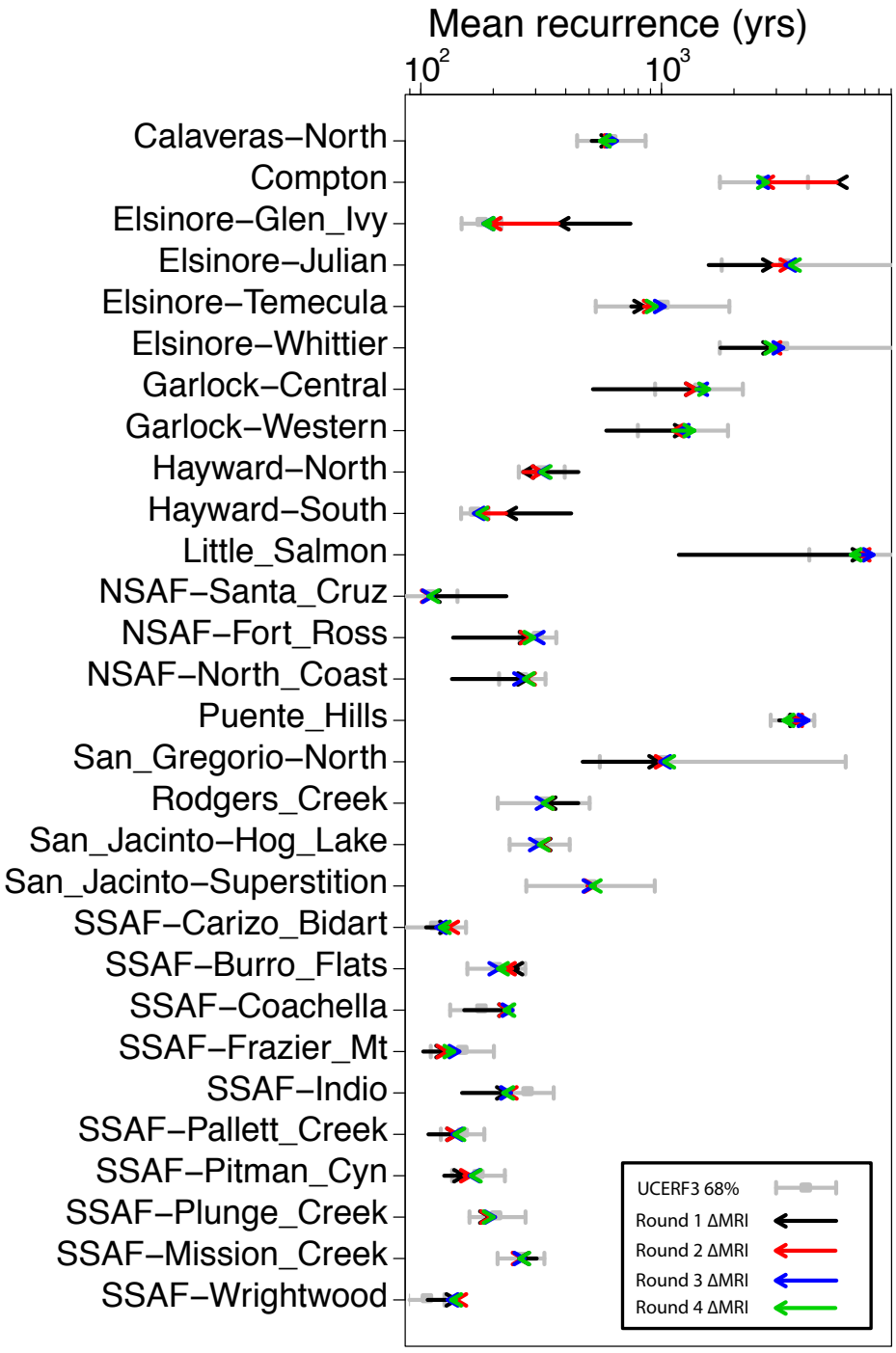

Figure 3.8: The change in the mean recurrence intervals for all sites over 4 tuning iterations are shown for the Under-Detection 'Paleoseismic' Catalog. The heads of the arrows point to the mean recurrence interval for each tuning iteration, while the tails of the arrows show the overall change from the last iteration. The 68% confidence intervals for the UCERF3 mean recurrence intervals are shown in grey. Note that these sites are in alphabetical order, not N-S or along-strike.

3.3.2 The Four Differently Tuned Catalogs

The four different catalogs were tuned based on different probabilities of detection and estimates of the mean recurrence intervals at the paleoseismic sites. The catalogs were designed to cover the full spectrum of event detectability (Under-, Complete-, and Over-detection) in order to investigate the strength of each of the possible explanations for the apparent hiatus in modern earthquakes. The effects of these different models of event detectability on both the distributions of events and the conditional probabilities are explored for each paleoseismic site in Section 3.4.

Catalogs were thinned during the tuning stage, as well as during the final processing stage (on the final, long catalog) according to the different probability of detection functions plotted in Figure 3.7. All catalogs were thinned during tuning based on the minimum rupture area requirement, but only two of them (the Under-detection catalogs) had events thinned out based on the minimum time between events or probability of detection because the model of detectability model assumed 100% detectability for the other two catalogs. Table 3.1 indicates the differences between these 4 catalogs for both the tuning and processing stages. It should be noted that some sites were more difficult to tune than others. It was particularly difficult to match the mean recurrence intervals for sites whose along-strike neighbors had significantly different mean recurrence intervals. Ironically, the Wrightwood Site on the Southern San Andreas Fault, which is considered to be one of the best-studied paleoseismic sites in California, could not be fit as well as many other sites. This is due to the fact that its mean recurrence interval is 41% - 64% lower than the Pallett Creek and Pitman Canyon sites on either side

of it. Even with large changes in the normal stress, it was difficult to get the site to rupture less often than the nearby sites along the same fault. However, the Wrightwood Site is still within the 68% confidence interval. The Alder Creek Site, on the Northern San Andreas Fault, has a reported mean recurrence interval that is 184% higher than the next site along strike. This could not be fit without imparting very large normal stresses (in excess of 200 MPa, over 100 MPa higher than the mean of the model), so it was not used in this study. Additionally, the Green Valley site turned out to be very difficult to fit, likely because that fault should have aseismic creep (which can be modeled with RSQSim but is not yet built into this fault model), so that site was also not used in this study. Overall, the remaining 29 sites are essentially over-fit, as the mean recurrence intervals are matched to well within the 68% confidence intervals.

The Under-Detection 'Paleoseismic' Catalog was thinned based on the minimum rupture area and time between events, as well as the UCERF3 probability of detection shown in blue on Figure 3.7 (UCERF3, Appendix I). For this model, the probability of detection increased with the amount of slip at a paleoseismic site. After thinning, it was tuned to match the UCERF3 paleoseismic recurrence intervals. The final, long catalog that was run with the tuned parameters was thinned once again, based on the same model of detectability. This thinned catalog represents a paleoseismically observable catalog (assuming the UCERF3 recurrence intervals are correct) as it contains only the events that would have been detected in paleoseismic trenches according to the UCERF3 paleoseismic detectability model. The Under-Detection 'Instrumental' catalog is the full version of the Under-Detection 'Paleoseismic' catalog. It was only thinned based on the

minimum event area (to avoid potential modeling biases) and all of the events that were 'not detected' based on the probability of detection or combined because they occurred too close in time in the 'Paleoseismic' version were kept in this 'Instrumental' version. This catalog represents the UCERF3 'Historic' or 'Instrumental' catalog, and all events that would have been detected by modern seismic instrumentation are present in the catalog. The Complete Detectability catalog was only thinned based on the minimum rupture area and a 100% probability of detection as a function of slip (i.e. events with any amount of slip were 'detected' and kept in the catalog) and tuned to match the UCERF3 paleoseismic recurrence intervals. This catalog assumed that the UCERF3 paleoseismic recurrence intervals were acurate and that the detectability was perfect. The last catalog was only thinned based on the minimum rupture area, and tuned to mean recurrence intervals that are 25% longer than the UCERF3 paleoseismic recurrence intervals. This 'Over-Detection' catalog assumed that a quarter of the events at the UCERF3 paleoseismic sites were spurious (i.e. 25% of trench events were not caused by earthquakes) and that event detectability was perfect. The increase in mean recurrence interval of 25% was chosen somewhat arbitrarily to investigate the effects of a significant increase in mean recurrence interval.

	Tuning	*Tuning*	*Tuning*	*Processing*	*Processing*
RSQSim Catalog	**Mean Recurrence Interval Goal**	**Combined Events with Minimum Δt**	**Thinned Using UCERF3 Probability of Detection**	**Combined Events with Minimum Δt**	**Thinned using UCERF3 Probability of Detection**
Under-detection 'Paleoseismic'	UCERF3	Yes	Yes	Yes	Yes
Under-detection 'Instrumental'	UCERF3	Yes	Yes	No	No
Complete Detection	UCERF3	No	No	No	No
Over-detection	125% UCERF3	No	No	No	No

Table 3.1: Details for the four simulated California catalogs to specify the differences between them in the tuning and processing steps. The catalogs tuned to the UCERF3 Mean Recurrence Interval Goal were tuned the match the UCERF3 mean recurrence intervals from Appendix H, while the catalog tuned to 125% of the UCERF3 Mean Recurrence Interval Goal, was tuned to mean recurrence intervals that were 25% longer than those from Appendix H. Combined Events with Minimum Δt are events that were too close in time (15 years or less) to be distinguished as separate events in a paleoseismic trench. Catalogs Thinned Using UCERF3 Probability of Detection were thinned using the blue UCERF3 Probability of Detection Function shown in Fig. 3.6, and the others were thinned using the black Complete Detection Function from Fig. 3.6.

3.4 Effect of Different Models of Detectability

3.4.1 Recurrence Interval and Magnitude Distributions

The different models of detectability had significant effects on the recurrence interval distributions and event probabilities, especially for the sites that were already past or near their mean recurrence intervals. The conditional probabilities of sites with mean recurrence intervals near the time since the most recent event are particularly sensitive to changes in both the modeled mean recurrence interval and Coefficient of Variation (COV) because the end time (T_{elapsed} + 30 years) used for the conditional probabilities extends well into the tail of the empirical recurrence distribution. When this

is the case A (from Equation 3.1) is very similar to B and the conditional probability is large.

Recurrence interval and magnitude distributions at the Rodgers Creek paleoseismic site are shown in Figure 3.9 for all four catalogs. This site was identified as a representative example of sites whose recurrence interval distributions and probabilities were most affected by the changes in detectability. The recurrence interval distributions are shown in red on the left side of Figure 3.9, with dt being the time (in years) between events. The UCERF3 mean recurrence interval (the tuning goal for the first three catalogs) is marked in blue at 325.3 years, the time since the most recent event detected in a trench at that site (303 years) is marked in green, and the RSQSim mean recurrence interval (calculated for each simulated catalog) is marked by a black dotted line. All three lines are very close together for the 'Paleoseismic' and Complete-detection catalogs because the UCERF3 MRI and T_{elapsed} are only 22 years apart, and those catalogs were both tuned to the UCERF3 mean recurrence interval and processed with the same probability of detection that they were tuned with. The thinning step in the processing stage shifted the recurrence distributions toward longer mean recurrence intervals and reduced the mean magnitudes of events. This was most apparent in comparisons between the 'Paleoseismic' and 'Instrumental' catalogs because they were tuned to the same mean recurrence interval, but the 'Paleoseismic' catalog was thinned substantially (based on slip) during the processing stage. The differences were primarily due to the removal (during thinning) of many of the small events in the 'Paleoseismic' catalog being combined because they often occurred closer in time than large events, or not being

'detected' because they had small amounts of slip. The difference in detectability between the 'Paleoseismic' catalog and the 'Instrumental' and Complete-Detection catalogs lengthened the simulated mean recurrence interval for the paleoseismic sites by almost twenty years. Additionally, there is a small but noticeable peak in the recurrence interval distribution of the 'Paleoseismic' catalog at ~650 years, which is roughly twice the mean recurrence interval for that site. This is due to some of the larger events with longer recurrence intervals being thinned out and leaving a much longer (in some cases doubled) recurrence interval between remaining events. These catalogs also tended to have smaller conditional probabilities. Ratios of the 30-year conditional probabilities for all sites from the 'Paleoseismic' catalog (which was thinned during processing of the long, tuned catalog prior to computation of probabilities) compared to the 'Instrumental' catalog (which was not thinned during processing) are shown in Figure 3.10. This ratio was less than one for most sites, indicating that the conditional probabilities were higher for the thinned sites. This is mainly because the recurrence times were longer and the COVs were larger for most of the thinned catalogs. The few sites that had probability ratios that were less than one were sites that already had very low conditional probabilities (<1%) and small COVs (0.15-0.25). In these cases, the change in COV due to thinning was very small (<0.04), so much of the variation is likely due to the randomness associated with thinning based on a probability that some events are not detected.

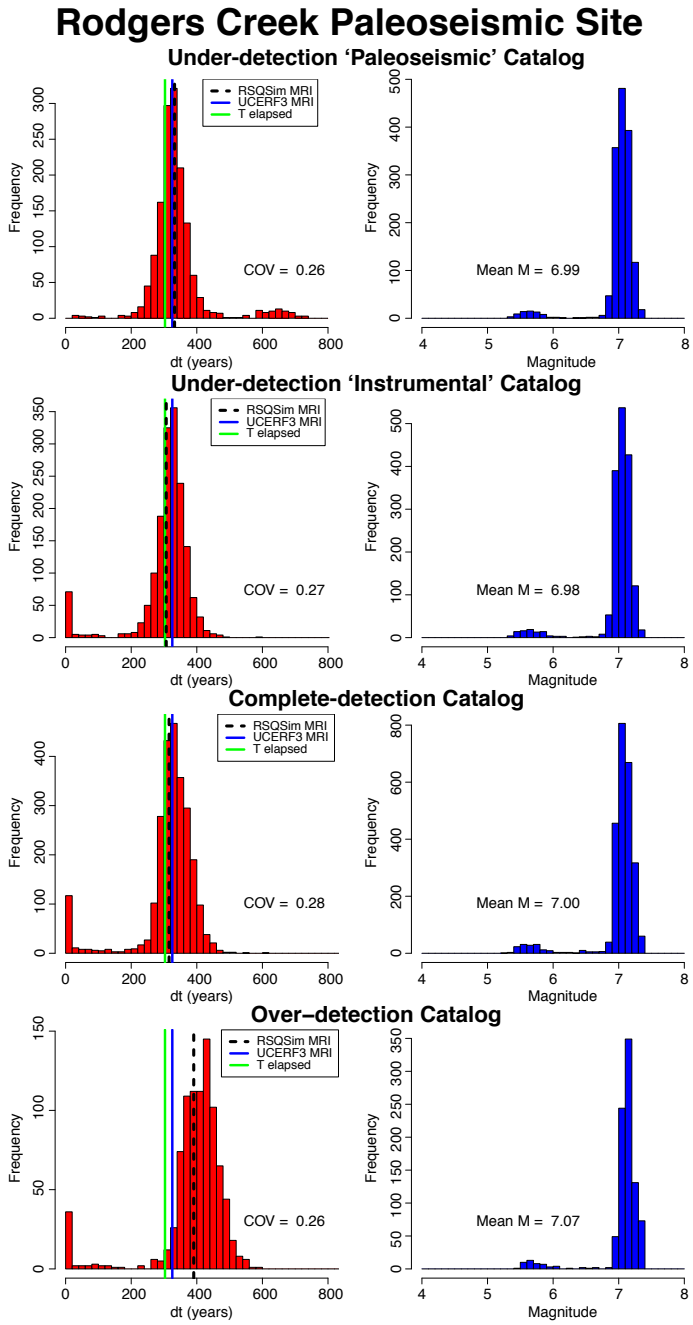

Figure 3.9: Recurrence interval (left) and magnitude (right) distributions from a representative paleoseismic site (Rodgers Creek) of all 4 catalogs for comparison of the affects of different probabilities of detection on the mean recurrence interval at this site. The elapsed time since the most recent event ($T_{elapsed}$) is indicated by a green line, the UCERF3 reported mean recurrence interval (UCERF3 MRI) is indicated by a blue line, and the modeled mean recurrence interval (RSQSim MRI) is indicated by a black, dashed line. The coefficient of variation of the recurrence intervals is listed for each catalog, as well at the mean magnitude of events.

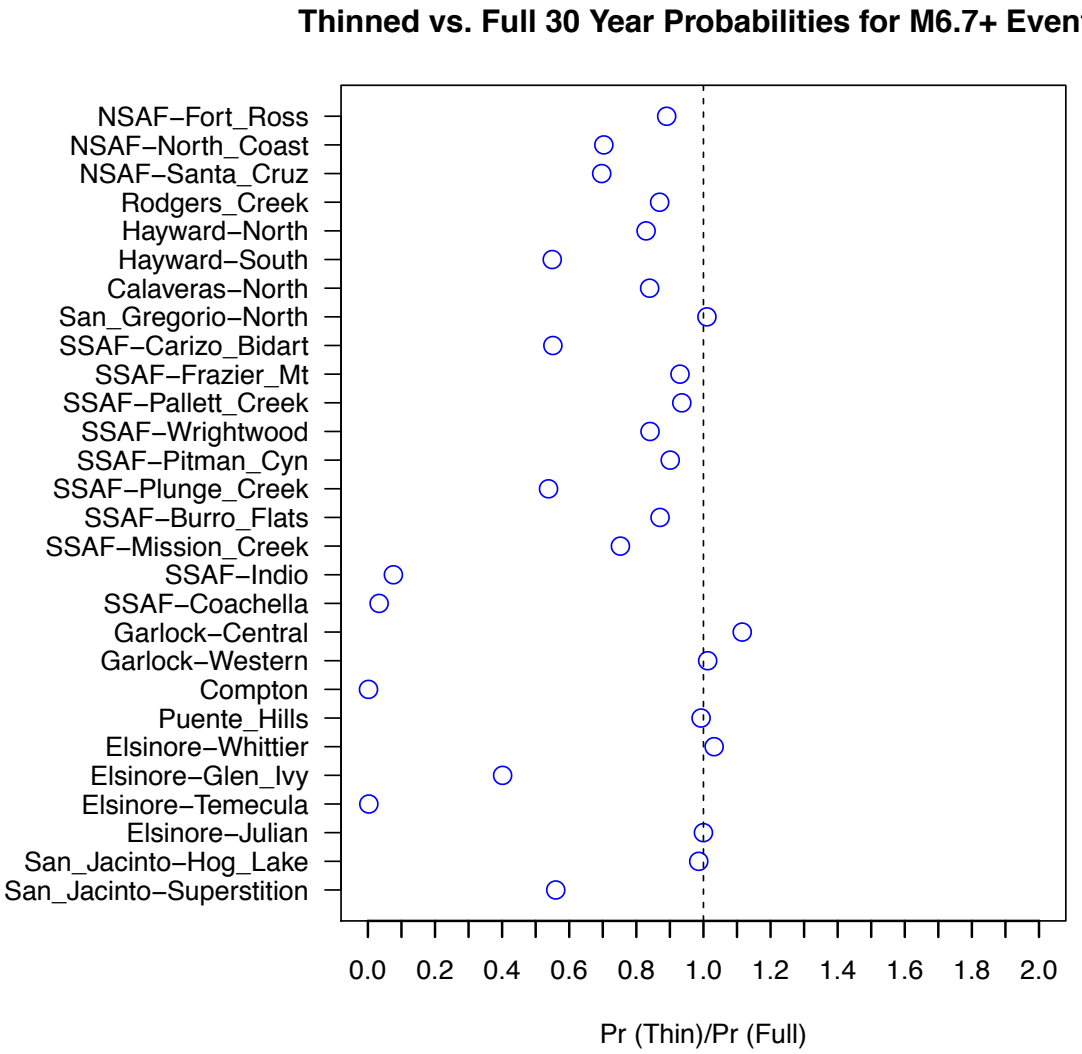

Figure 3.10: Ratios of the 30-year conditional probabilities of M6.7+ events for the Under-detection 'Paleoseismic' catalog (which is thinned based on which events would be paleoseismically detectable) vs. the Under-detection 'Instrumental' catalog (which is not thinned under the assumption that all events would be instrumentally detectable). Note that thinning the catalog primarily reduced the conditional probabilities.

3.4.2 Conditional Probabilities

Many paleoseismic sites are past or very close to their MRI, which yields exceptionally high conditional probabilities from RSQSim catalogs. This is due to both the short recurrence intervals at these sites and the relatively small coefficients of variation that the simulations generate. The 30-year conditional probabilities for large events (M6.7+) at each paleoseismic site are shown in Figure 3.11 and the coefficients of variation for the recurrence times at each site are shown in Figure 3.12. The probabilities at the Southern San Andreas sites, as well as several Northern California sites, are particularly high because many of these sites are past their mean recurrence intervals. However, valid conditional probabilities could not be calculated for three Southern San Andreas sites (Carizo, Indio, and Coachella) because so few events had recurrence intervals between the time since the most recent event and the end of the recurrence distribution (i.e. there were only a few events used to calculate the value B from conditional probability Equation 3.1). The Over-detection model, which had the longest recurrence intervals (25% longer), generated what seem to be the most reasonable conditional probabilities (i.e. the lowest). This model could explain why we don't see as many earthquakes now as would be inferred from paleoseismology. The conditional probabilities for the 'Paleoseismic' catalog tended to be the next lowest, followed by the Complete-detection and 'Instrumental' catalogs. The 'Paleoseismic' catalog has consistently lower probabilities than the 'Instrumental' catalog, implying that reduced detectability, or at least additional thinning, may reduce probability estimates.

The coefficients of variation of the recurrence distributions in the current model are at the lower limit thought to exist in nature of 0.4 to 0.7 for events occurring at a point on a fault, such as a paleoseismic site (Field et al., 2014). Overall, the catalogs with the worst detectability, i.e. the ones that were thinned the most, had the largest COVs. Excluding the Over-detection catalog (because it was tuned to longer recurrence intervals) the catalogs with the lowest conditional probabilities also tended to have the highest COVs, suggesting that greater variation in the recurrence times for events at each site (and thus larger COVs) would result in decreased conditional probabilities overall.

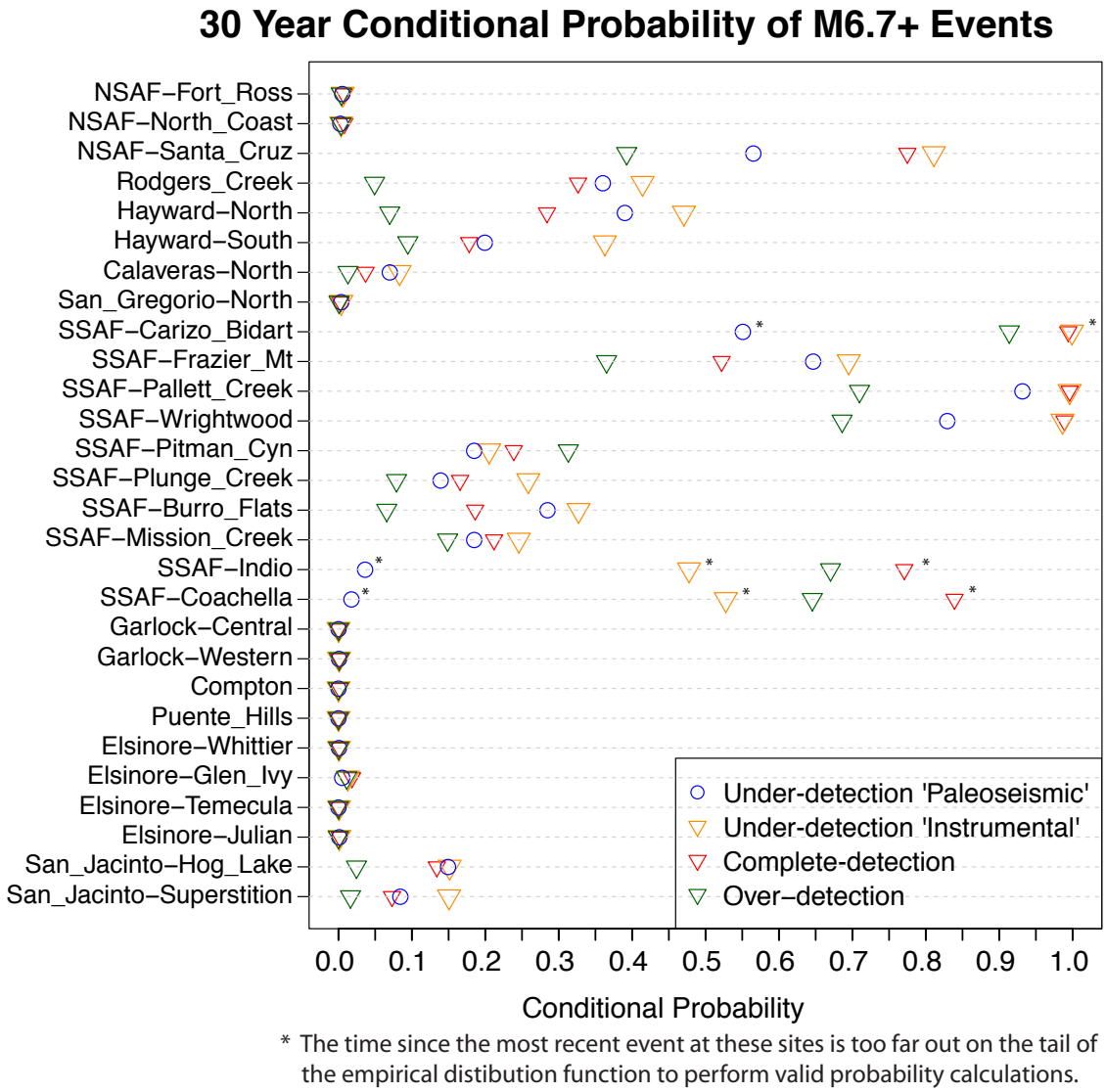

Figure 3.11: 30-year Conditional Probabilities of magnitude 6.7+ events for all four of the differently tuned, simulated catalogs. Overall, the probabilities are extremely high, except for the Over-detection catalog, which was tuned to 25% longer mean recurrence intervals, suggesting that the reported paleoseismic mean recurrence interval are too high.

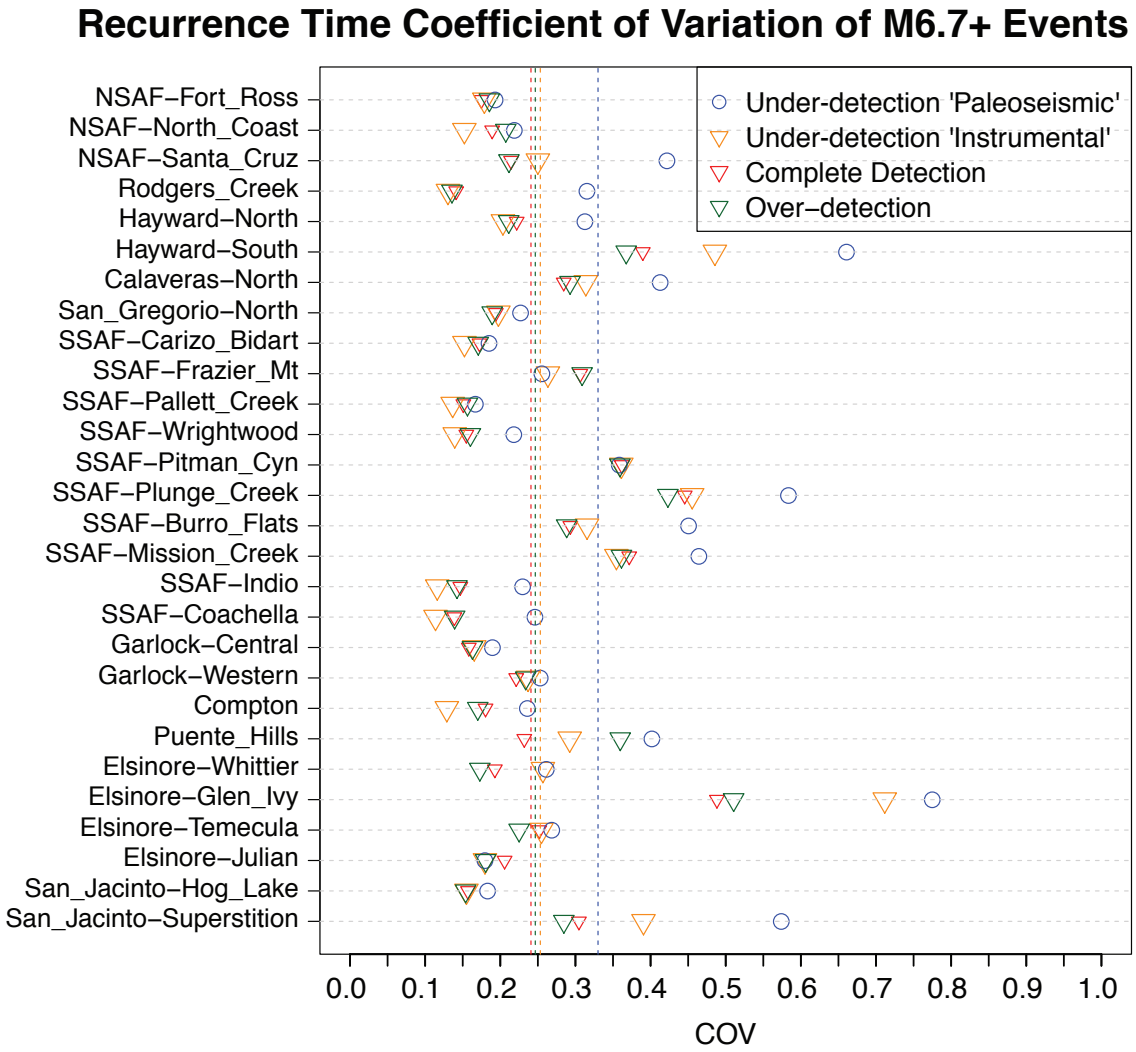

Figure 3.12: Coefficients of variation of the recurrence intervals for each of the paleoseismic sites, with the means marked by the vertical dashed lines. Many of the COVs are very small compared to the values thought to exist in nature of 0.4 to 0.7 for events occurring at a point on a fault, such as a paleoseismic site (Field et al., 2014).

3.4.3 Sites Past Their Mean Recurrence Interval

There are currently eleven paleoseismic sites past their UCERF3 mean recurrence interval. Based on the mean recurrence intervals and times since the last event at each site, David Jackson has (SCEC 2015, Abstract #081; personal communication) determined that the probability of none of these sites rupturing in the last 100+ years is

<1%. This strongly suggests that there is a misfit between the inferred paleo- and modern-event rates. However, this metric is not directly comparable to the observation that more than one-third of the sites are past their mean recurrence time. The RSQSim catalogs are long and detailed enough to investigate the number of paleoseismic sites that are past the UCERF3 mean recurrence intervals at any given time. Figure 3.13 is a representative 10,000-year snapshot from the Under-detection 'Paleoseismic' catalog (which was tuned to be the closest match to the UCERF3 data) showing the variation in time of the number of paleoseismic sites that are past their UCERF3 mean recurrence interval in a given year. This data was compiled by performing a yearly count of how many sites had a Tmre that was longer than their reported MRI. There are usually only a few sites past the UCERF3 mean recurrence interval and there are never more than fourteen sites past the UCERF3 mean recurrence interval during the entire 500,000-year catalog. In fact, the average number of sites past the UCERF3 mean recurrence intervals over time, for the catalogs that were tuned to the UCERF3 mean recurrence intervals, is 3.2.

Figure 3.14 illustrates the probabilities of any number of paleoseismic sites being past the UCERF3 mean recurrence intervals at any time (top), as well as at times after there had been both an 1857- and 1906-like event in a 50-year period (bottom) based on the counts performed for Figure 3.13. The probabilities of eleven sites being past their UCERF3 mean recurrence intervals in the RSQSim catalogs that were tuned to the UCERF3 mean recurrence intervals (Under-detection 'Paleoseismic', Under-detection 'Instrumental', and Complete Detection) are all under 0.1% (0.07%, 0.003%, and 0.06%

respectively). However, the Over-detection catalog, which was tuned to 25% higher mean recurrence intervals than the reported paleoseismic means, yields a more reasonable 17.5% probability of eleven sites being past the UCERF3 mean recurrence intervals.

While one might expect that the occurrence of two large events within a short period of time (as in the case of the 1857 and 1906 earthquakes) would reduce event rates for a while and thus allow more sites to reach their mean recurrence intervals, this metric actually makes the problem seem worse. The probabilities for all numbers of sites being past their recurrence intervals are decreased, simply because nine of the paleoseismic sites had (by definition of 1857- and 1906-like events) recently ruptured and thus couldn't be near their mean recurrence intervals.

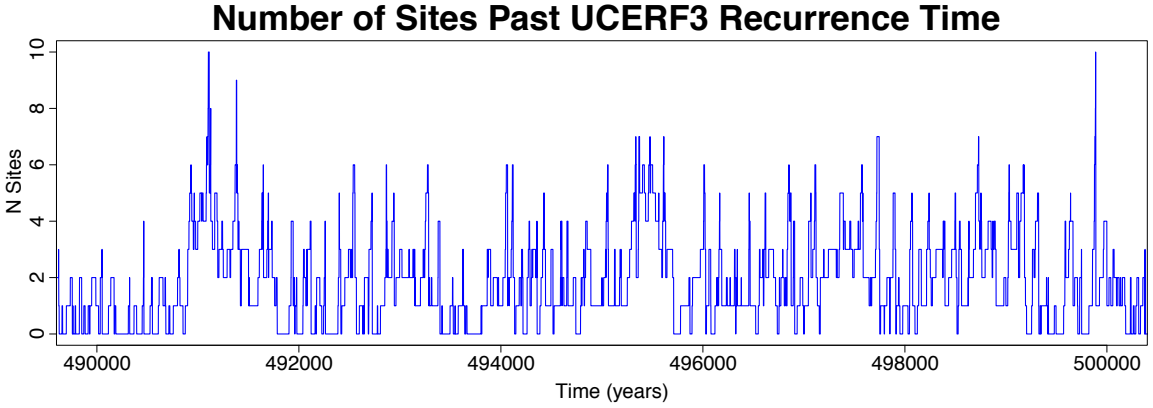

Figure 3.13: 10,000 year snapshot from the Under-detection 'Paleoseismic' catalog, showing the number (N) of paleoseismic sites that are past their UCERF3 mean recurrence interval in a given year. N is quite variable, but at or below 2 sites 85% of the time and never reaches 11 (the current number for California, at least during this time period indicating that such an occurrence rare.

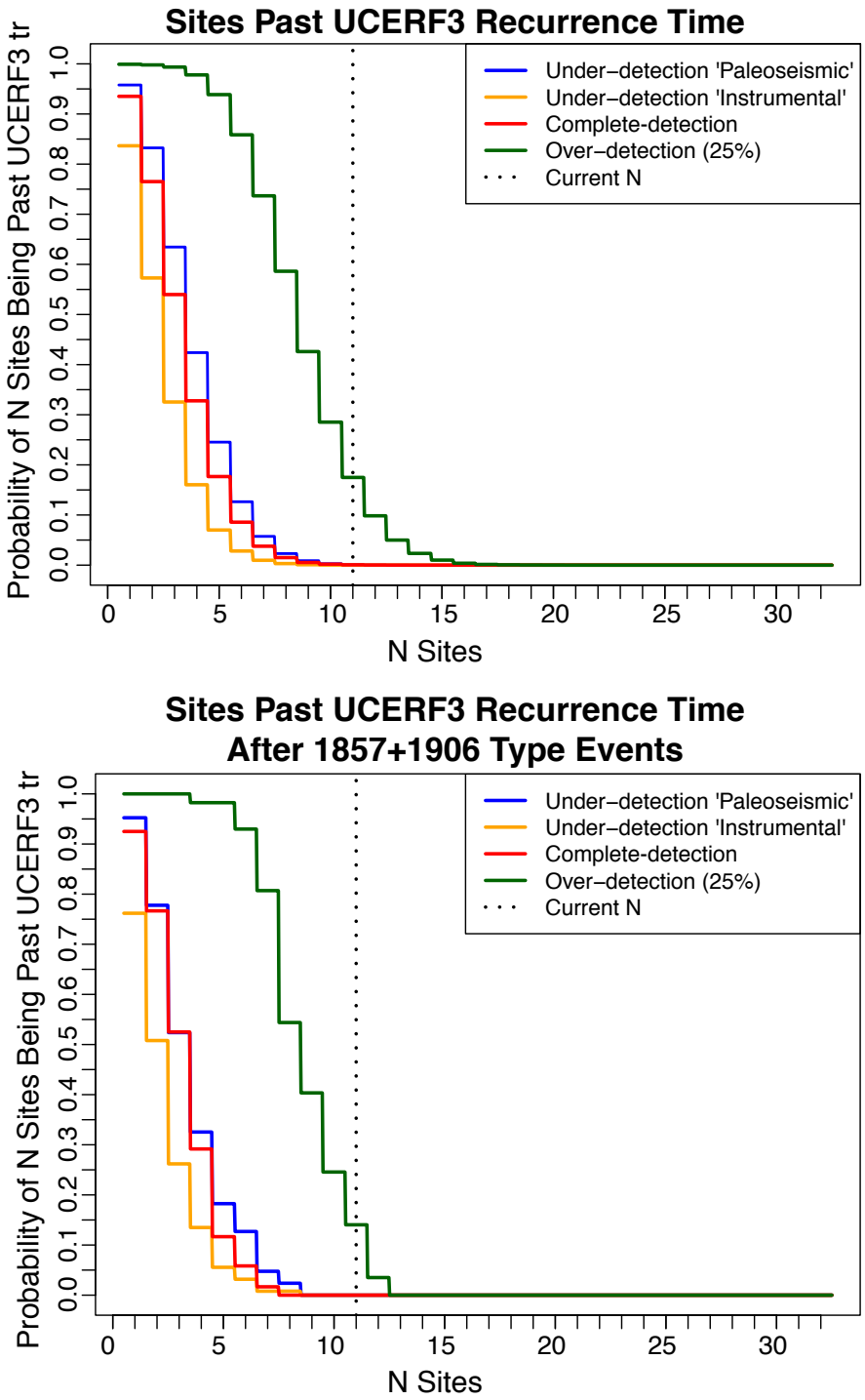

Figure 3.14: The probability that N sites will be past the UCERF3 mean recurrence interval at any given time is shown above for all events in all four simulated catalogs (left), as well as after the occurrence of 1857 + 1906 type events have occurred (right). The probability of the current N=11 is <0.1% for all catalogs, excluding the Over-detection catalog, suggesting that the current situation is highly unlikely for the reported mean recurrence times.

3.5 Conclusions and Future Model Improvements

Currently in California, the number of paleoseismic sites that are at or past their UCERF3 mean recurrence interval is very high and it is extremely unlikely (<0.1% in our simulations) that it can be attributed to coincidence or good luck. It is possible that California is in a period of lower event rates due to stress shadowing from the 1857 and 1906 events, but the current simulator data does not support this. Given the importance of resolving the possible discrepancies between expected and observed rates of damaging earthquakes since 1906, further in-depth study of this topic is warranted. There are several possibilities that can be explored. Reducing some of the long-term slip rates in the model may yield more reasonable probabilities and make it easier to tune the model to some of the difficult paleoseismic sites such as the Wrightwood paleoseismic site. Based on the results shown in Figure 3.13, some of the recurrence intervals from paleoseismic trenches may be too short, by as much as 25% or more. This could be because they include some small events, as there is a small probability that they would be detected in trenches, or some features in the trenches could be due to causes other than primary earthquake ruptures. Over-detection at paleoseismic sites could add spurious events to the catalogs, thus reducing the long-term mean recurrence intervals. If our models are correct, and the paleoseismic mean recurrence intervals are artificially elevated, this would suggest that the seismic hazard near some of these site is lower than previously thought. Overall, the conditional probabilities for large events at paleoseismic sites in our RSQSim catalogs are suspiciously high however. This may be a modeling issue related to the narrow recurrence interval distributions (and thus small COVs). Adding off-fault

seismicity and deep creeping sections that undergo afterslip following large earthquakes may add more variation to the event times and increase the COVs. However, the most reasonable conditional probabilities that are closer to those from the UCERF3 report came from the catalog that was tuned to longer recurrence intervals than those reported by UCERF3. Results also show that it is extremely unlikely that we could be in an interval with so many sites past their UCERF3 mean recurrence interval, so the problem cannot be explained entirely by potential modeling issues.

3.6 References

Cornell C.A. (1968). Engineering Seismic Risk Analysis. "Bulletin of the Seismological Society of America", 58(5):1583-1606.

Field, E. H., T. E. Dawson, K. R. Felzer, A. D. Frankel, V. Gupta, T. H. Jordan, T. Parsons, M. D. Petersen, R. S. Stein, R. J. Weldon II, and C. J. Wills (2008). The Uniform California Earthquake Rupture Forecast, Version2 (UCERF 2), USGS Open File Report 2007-1437.

Field, E. H., Arrowsmith, R. J., Biasi, G. P., Bird, P., Dawson, T. E., Felzer, K. R., ... & Zeng, Y. (2014). Uniform California Earthquake Rupture Forecast, Version 3 (UCERF3)—The Time-Independent Model. *Bulletin of the Seismological Society of America*, *104*(3), 1122-1180.

Harris, R. A., & Simpson, R. W. (1998). Suppression of large earthquakes by stress shadows: A comparison of Coulomb and rate‐and‐state failure. *Journal of Geophysical Research: Solid Earth (1978–2012)*, *103*(B10), 24439-24451.

Jackson, D.D., Richards-Dinger, K.B. (2015). Recurrence Implications of California Paleo-event Haitus. Poster 081. Southern California Earthquake Center Annual Meeting, Palm Springs, California, 12-16 September.

McCalpin, J. P. (Ed.). (2009). *Paleoseismology* (Vol. 95). Academic press.

Simpson, R. W. and Reasenberg, P. A., in US Geological Survey Professional Paper 1550–F (ed. Simpson, R. W.), 1994.

Chapter 4

Earthquake Clustering in Simulated Catalogs

4.1 Abstract

Clusters of large earthquakes, such as the Joshua Tree–Landers–Hector Mine sequence in California in the 1990s indicate that the occurrence of one large earthquake may increase the probability of additional large earthquakes nearby. However, the California earthquake catalog spans too short of a duration to define the statistics of large-event clusters. The 3D boundary element code RSQSim generates simulated California earthquake catalogs with millions of M≥4 events, spanning from tens of thousands up to a million years, which allows for the investigation of both the statistics and physical characteristics of large-event clusters. We use RSQSim to investigate the characteristics and probabilities of large-event clusters in California catalogs, as well as catalogs from simpler, idealized fault systems. The RSQSim simulations incorporate rate-state fault constitutive properties, and the simulated catalogs include foreshocks, aftershocks, and occasional large-event clusters. The vast majority of large-event clusters are event pairs consisting of only two events with occasional three- and four-event clusters. For the primary purposes of this study a large-event cluster is defined as two or more M≥7 events that occur within four years and one rupture length apart. Rates of large-event clusters observed in RSQSim catalogs are higher than the expected rates for a truly random

process, suggesting that secondary events in large-event clusters are triggered by the primary events. Additionally, rates of clustering are affected by several factors including rate- and state-dependent constitutive parameters, fault complexity, mainshock magnitude, and the inter-event time of the clustered events. Overall, large-event clustering, which is typically not considered in long-term earthquake forecasts, represents a significant probability gain for large earthquakes.

4.2 Introduction

It is well known that large earthquakes occasionally occur in space-time clusters, in which two or more large events occur within a small area, i.e. tens of kilometers, and a short period of time, i.e. a few years (Kagan & Jackson, 1991; Dieterich, 1994). Some notable large-event clusters include: recurring event pairs along the Nankai Trough off the coast of Japan (M8.4, Dec. 23, 1854; M8.4, Dec. 24, 1854; M8.1, Dec. 7, 1944; M8.1, Dec. 24, 1946), the New Madrid Sequence of M≥7.5 events in the Central United States (Dec. 16, 1811; Jan. 23, 1812; Feb. 7, 1812), and the series of events in Southern California including the Joshua Tree (M6.1, Apr. 23, 1992), Landers (M7.3, Jun. 28, 1992), Big Bear (M6.5, Jun. 28, 1992), and Hector Mine (M7.1, Oct. 16, 1999) earthquakes. Studies have shown that the short-term probability of an additional earthquake with a significant moment release, i.e. M≥7, increases in the general area of a mainshock (Kagan and Jackson, 1999; Wang et al., 2010). Although clusters of large events are relatively rare, understanding these large-event clusters has important implications for estimating earthquake hazard. This is especially important for disaster

planning and relief efforts where a second or third event striking already damaged and unstable structures can be devastating, and interrupt rescue and recovery efforts. However, current methods for producing long-term earthquake forecasts do not consider clustering and thus systematically underestimate earthquake hazard (UCERF3, Field et al., 2014).

Although clusters of large events occur throughout the world, instrumental earthquake catalogs are too short in duration to define the statistics of large-event clusters in specific regions. Historical records give information about the timing and intensity of large earthquakes for some regions, but the details of these events are limited. Japanese records, for example, indicate that large-event clusters have been occurring along the Nankai Trough for at least a thousand years (Ando, 1975) but only one of these, a large event pair in 1944 and 1946, has occurred since modern seismological instruments were developed. Most instrumental earthquake catalogues span just over a hundred years, and although paleoseismic studies extend this time scale for some fault systems, such studies generally lack the temporal resolution to distinguish short-term clustering (Scharer et al., 2011). While paleoseismic and dating techniques for estimating earthquake recurrence are improving, errors on age estimates are still on the order of 20+ years and these studies depend heavily on the preservation of ruptures and availability of datable material (Scharer et al., 2011). Furthermore, paleoseismic records do not have well-constrained, detailed information about magnitude, location, rupture extent, or aftershock and foreshock sequences. The limitations on observing earthquake occurrence leave many unanswered questions. For example, are the recurring earthquake pairs along the Nankai

Trough (the data for which are further limited by the impossibility of paleoseismic studies offshore) due to random chance, or consistent physical features of the fault system? What controls the timing, location, and extent of these ruptures and are there indicators that another section of the fault system is going to rupture shortly? There are simply not enough combined temporal and spatial data within a specific region to address these questions on a purely statistical basis. Stacking worldwide catalogs has revealed global statistical characteristics of space-time clusters (Kagan and Jackson, 1999), but is insufficient to reveal the properties that may act as local controls of clustering. These properties may include the fault geometry, long-term slip rate, and fault constitutive properties.

Refining earthquake hazard analyses and forecasting models has been one of the main goals of the earthquake physics community for decades. Previous studies have focused primarily on statistical analysis of catalog data and stochastic models to calculate earthquake probabilities. Physical models based on Coulomb stress analysis and rate- and state-friction have also been used to determine earthquake probabilities (King et al., 1994; Stein et al., 1997). The earliest characterizations of earthquake occurrence and clustering were based on statistical models of relatively limited catalogs. Omori (1894) described the short-term clustering of earthquakes based on the rate of occurrence of aftershocks. Later studies (Reasenberg, 1985; Ogata 1988; Kagan & Knopoff, 1987) have focused extensively on models of Poisson processes to identify random, independent events and thus separate the clustered events. Kagan and Jackson (1991,1999) demonstrated that large as well as small earthquakes cluster in both time and space. All of

these studies however, are based on less than one hundred years of instrumental seismic data (much less in most cases), which is not enough time to get a reliable measure of the recurrence rates of large earthquakes. This time scale is particularly problematic if the recurrence rates vary with time.

Stochastic earthquake clustering models have been used to estimate short-term earthquake probabilities. Ogata (1988) first introduced the Epidemic-Type Aftershock Sequence Model (ETAS) and it has been used to model the time, space, and magnitude dependence of seismicity throughout the world. The Short-Term Earthquake Probability Model (STEP), a one-generational ETAS model in which aftershocks do not produce their own aftershock sequences, has been used by the USGS to estimate short-term earthquake probabilities. While these methods are based on empirical aftershock and magnitude frequency relations and have done well to replicate some aspects of seismicity, they do not explain the physical processes behind space-time clustering nor are they capable of characterizing space-time clustering on specific faults.

Physical models based on rate- and state-dependent friction have become increasingly prevalent for studying earthquake clustering (Dieterich, 1994; Dieterich & Kilgore, 1996; Toda and Stein, 2002; Hainzl et al., 2003; Dieterich & Richards-Dinger, 2010). Dieterich and Richards-Dinger (2010) used RSQSim to study the effect of fault complexity on temporal clustering of large events in simple fault systems consisting of planar faults, fractally-rough faults, and fractally-segmented faults. That study highlighted the effects of rate- and state–dependent friction on earthquake recurrence, but only for highly simplified fault systems. Richards-Dinger and Dieterich (2012) studied

the spatial and temporal clustering of earthquakes in a much more realistic fault system (the UCERF2 fault model described in Chapter 1) and found that events that are close in time are close in space as well, indicating a causal relationship between clustered events.

This study implements the 3D boundary element code RSQSim with California fault models, to investigate clusters of large events. RSQSim has been successful in producing realistic earthquake catalogs that compare well to California catalogs and are long enough to provide useful clustering statistics (Richards-Dinger and Dieterich, 2012). It allows large-scale simulations of fault system seismicity, which produce clustered seismicity in the form of foreshocks, aftershocks, and occasional large-event clusters. Fault system simulations generate Gutenberg-Richter type magnitude frequency distributions and Omori-like aftershock distributions (Dieterich and Richards-Dinger, 2010). RSQSim generates synthetic catalogs with hundreds of thousands to millions of $M \geq 4$ events, spanning from tens of thousands, up to a million years, which allows for the investigation of the characteristics of long-term earthquake probabilities. Additionally, the space-time statistics of simulated California catalogs can be used to identify clustered and non-clustered populations of events and indicate the presence of seismicity rate changes, which are discussed in Section 4.4.

Dieterich (1994) proposed that rate- and state-dependent friction is the underlying cause for earthquake clustering. Unlike simple Coulomb friction, which involves effectively instantaneous nucleation of unstable slip at a fixed stress threshold, rate- and state-dependent friction involves time-dependent nucleation over a range of stress conditions. The time delay associated with the nucleation phase depends on the

magnitude of the stress perturbation. Consequently, aftershocks as well as the secondary events in large-event clusters are due to delayed failure from the stress perturbation applied at the time of the first event in large-event clusters.

4.3 Fault Models and Simulated California Catalogs

Data from two different California fault models are presented both here and in Chapter 5. The primary fault model used to test the effects of various fault properties on clustering statistics is the UCERF2 fault model described in Chapter 1 (Figure 1.2). Several catalogs were generated with this fault model in order to compare the effects of varying the rate- and state-dependent constitutive parameter *a*, from Equation 1.1 (Section 4.5.1). These catalogs span tens of thousands of years and contain a few hundred thousand M4.9-M7.9 events. Data from the UCERF3 'Paleoseismic' Catalog described in Chapter 3, which was produced using the UCERF3 fault model (also described in Chapter 3) is presented as well. This catalog spans over 500,000 years with over 74 million ~M4 to M7.9 events.

The original California fault model (UCERF2) had larger (3x3 km vs. 1x1 km) and fewer (12,500 vs. 126,000) fault elements. Those simulations were designed to be significantly less computationally expensive than the later, UCERF3-based models, so they could be run on computers in the UCR Earth Sciences Department. This enabled testing the effects of system-wide parameters such as the rate- and state-dependent friction constitutive parameter *a*, which required running several simulations. The new UCERF3 fault model was preferred for studying foreshock and aftershock statistics

because the smaller elements reduced the minimum earthquake magnitude to roughly M4 from nearly M5, which greatly increased the numbers of small events, and thus the number of foreshocks and aftershocks. The magnitude difference is due to the fact that the minimum magnitude in RSQSim simulations is limited by the size of fault elements because the smallest event that can occur is a single element event. However, decreasing the element size to 1-km^2, and thus increasing the number of elements by more than a factor of ten, necessitated running the simulations on a supercomputer.

Data from six different RSQSim catalogs are presented in this chapter. The two main catalogs are the million-year UCERF2 catalog and the half-million-year UCERF3 catalog. The main difference between these catalogs is the number and minimum size of the small earthquakes related to the different element sizes used. The rate- and state-constitutive parameters in both cases were kept the same as those presented in Chapter 1 (μ_0=0.6, a=0.01, b=0.014, and D_C=10^{-5}m). Data from four shorter duration catalogs (~50,000 years) from the UCERF2 fault model are also presented in Section 4.5, in which the rate- and state- constitutive parameter a was varied. Additionally, several completely synthetic Poisson catalogs, based entirely on empirical earthquake statistics without clustering, are presented in Section 4.4 for comparison with the physics-based RSQSim catalogs and the instrumental California catalog. The instrumental catalog used in this study is the ANSS Catalog for California from 1911 to 2015.

4.4 Space-Time Statistics

Figures 4.1 and 4.2 illustrate the inter-event time probability density distributions and inter-event distance vs. inter-event time distributions respectively, from the instrumental California catalog (ANSS, 1911-2015) and the simulated California catalog (UCERF2) for three different magnitude ranges. In these plots the inter-event time of two earthquakes is defined as the elapsed time between the occurrences of pairs of successive events within a specified magnitude range in a catalog, regardless of where they occur in the fault system. The inter-event distance is simply the spatial distance between the events forming each pair.

For comparative purposes, Figure 4.3 similarly shows inter-event time and inter-event distance plots from four synthetic Poisson catalogs that have neither spatial nor temporal clustering. The Poisson catalogs were created by randomly drawing event times and locations from uniform distributions, and event magnitudes from a Gutenberg-Richter distribution. To simulate a non-stationary Poisson process (colored curves in Figure 4.3), Poisson catalogs with different event rates and durations were produced and these catalogs were spliced together to create longer catalogs with non-stationary event rates.

The inter-event times for a stationary Poisson process have an exponential density distribution:

$$\rho_{Poison} = \lambda e^{-\lambda t}, \tag{4.1}$$

where $\boldsymbol{\lambda}$ is the event rate, and t is the inter-event time. The gray line in Figure 4.3 is the probability density distribution for a stationary Poisson catalog, which follows the

exponential distribution given by Equation 4.1. The colored lines give the density distributions for the combined non-stationary Poisson distributions. The inter-event time statistics of the simulated California catalogs can be fit by a combination of two different functions: the exponential function of Equation 4.1 (shown in blue in Figure 4.1), which is characteristic of a stationary Poisson process, and a power-law function (shown in red in Figure 4.1):

$$\rho_{Power} = A\tau^{-q}, \tag{4.2}$$

where τ is the inter-event time, A is a constant, and q is the slope of the inter-event time distribution. On a log-log plot the slope of this distribution, given by q, is closely related to the power law slope in the Omori aftershock decay law. The modified Omori's law (Equation 4.3, from Utsu, 1961) gives the rate of aftershocks with time after a mainshock:

$$n(t) = \frac{k}{(c+t)^p}\ , \tag{4.3}$$

Where k and c are constants that vary for different aftershock sequences, t is the time, and p is the slope or decay rate, which is typically about -1 but varies between -0.7 and -1.5. The relationship between p and q is:

$$q = 2 - (1/p), \tag{4.4}$$

where q is the slope of the inter-event time distribution and p is the aftershock decay rate (Ustu, 1995 – from Shenshu, 1959 in Japanese). The events that fit a power-law are the clustered events and those that fit an exponential distribution are the random, unrelated events. The truly unrelated events tend to have longer inter-event times and distances than the clustered events.

The ‘tails’ of the inter-event time distributions (i.e. the ends of the distributions at longer inter-event times – circled in red on Figure 4.1) deviate from the expected Poisson curves (blue curves in Figure 4.1). These tails can be explained by a non-stationary Poisson process where the rates vary with time as shown in Figure 4.3. The ‘tails’ move outward, away from the expected curve (that of a stationary Poisson process) as the event rate varies and the difference increases for greater changes in event rate. This is observed in both the simulated California catalogs and the instrumental California catalog. The rate variation in these catalogs, which are not primarily random and therefore not purely Poissonian, is likely due to periods of partial quiescence after the occurrence of large events.

The distinction between the power-law and exponential distributions becomes more apparent for larger magnitude earthquakes, as illustrated in Figure 4.1. Additionally, the transition between the power-law and Poisson distributions of large events ($M \geq 7$) occurs at inter-event times of a few to several years, indicating that the time between large events in large-event clusters is likely to be similar because inter-event times that are longer than this fall into the Poissonian part of the inter-event time distribution. According to Stein et al. (1997) successive $M \geq 7$ events on the North Anatolian fault have inter-event times that range from 3 months to 30 years, indicating that the time periods between clustered events in nature (at least in this location) are similar to those of simulated catalogs.

The space-time statistics of the California catalog and the simulated California catalogs are separated into different magnitude ranges (Figures 4.2 and 4.3). While there

are many fewer events in the instrumental California catalog compared to the RSQSim California catalog, the two catalogs show the same general characteristics for the smaller events, where there are enough data for a true comparison. There are not enough large events in the instrumental California catalog to study the statistics of large-event clusters because the catalog spans just over one hundred years and contains only nine $M \geq 7$ events. However, the broad agreement of the simulations with observations at lower magnitudes suggests that the simulations can provide a physics-based extrapolation of clustering statistics to large magnitudes where data are not available.

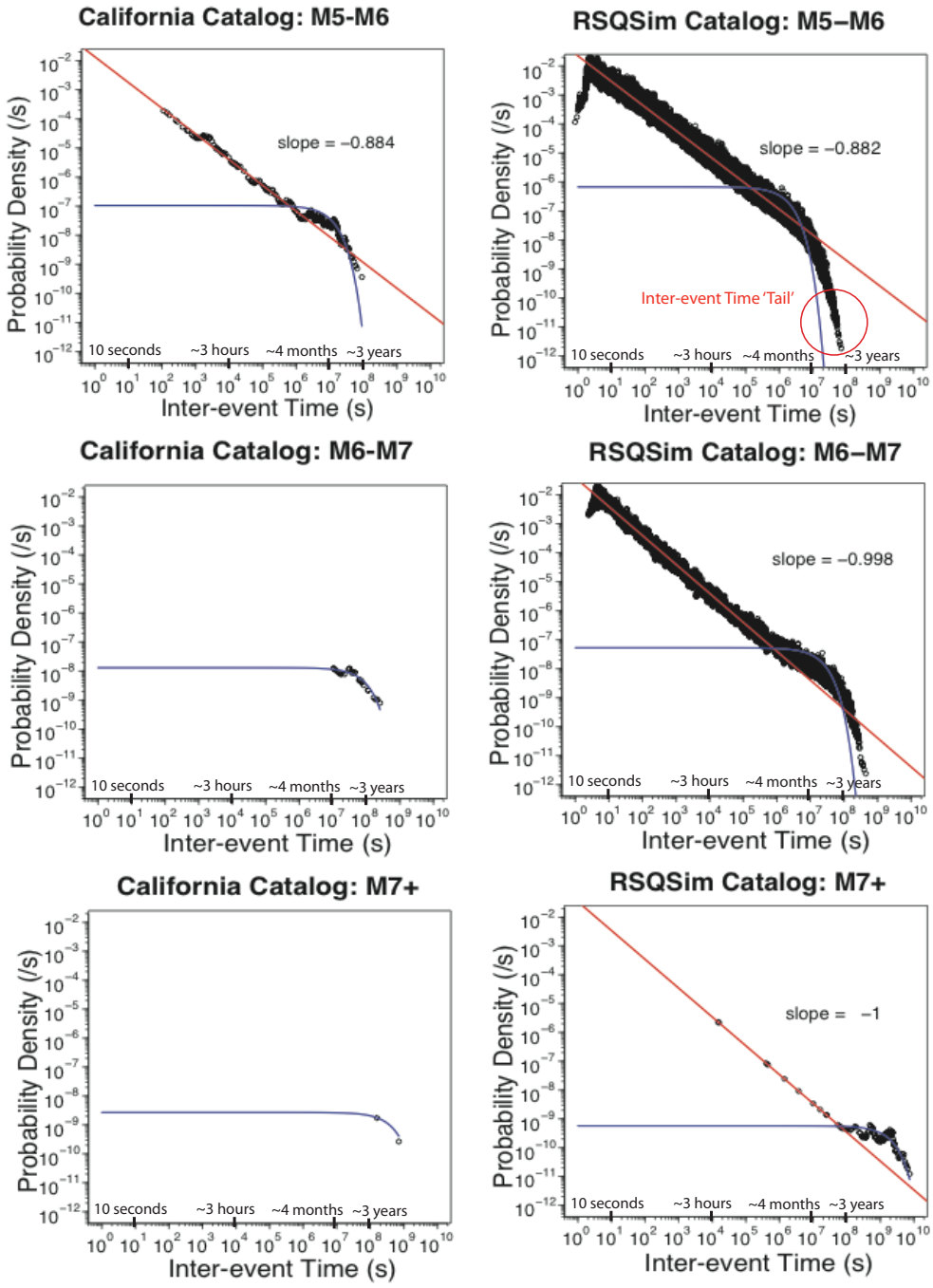

Figure 4.1: Inter-event time probability distributions for the ANSS California catalog from 1911-2010 (left) and the RSQSim UCERF2 catalog (right) for magnitude ranges M5-M6 (top), M6-M7 (middle), and M7+ (bottom). Red lines indicate a power law fit to the data (representing the clustered events) and blue lines indicate an exponential (Poisson) fit to the data (representing the random, unrelated events). Where there are enough data, the California catalogs show the same general trends as the simulated All California catalog. The red circle indicates where the 'tail' of the inter-event time distribution deviates from the expected exponential fit. This is due to variation in the event rate within the catalog as shown in Figure 4.2.

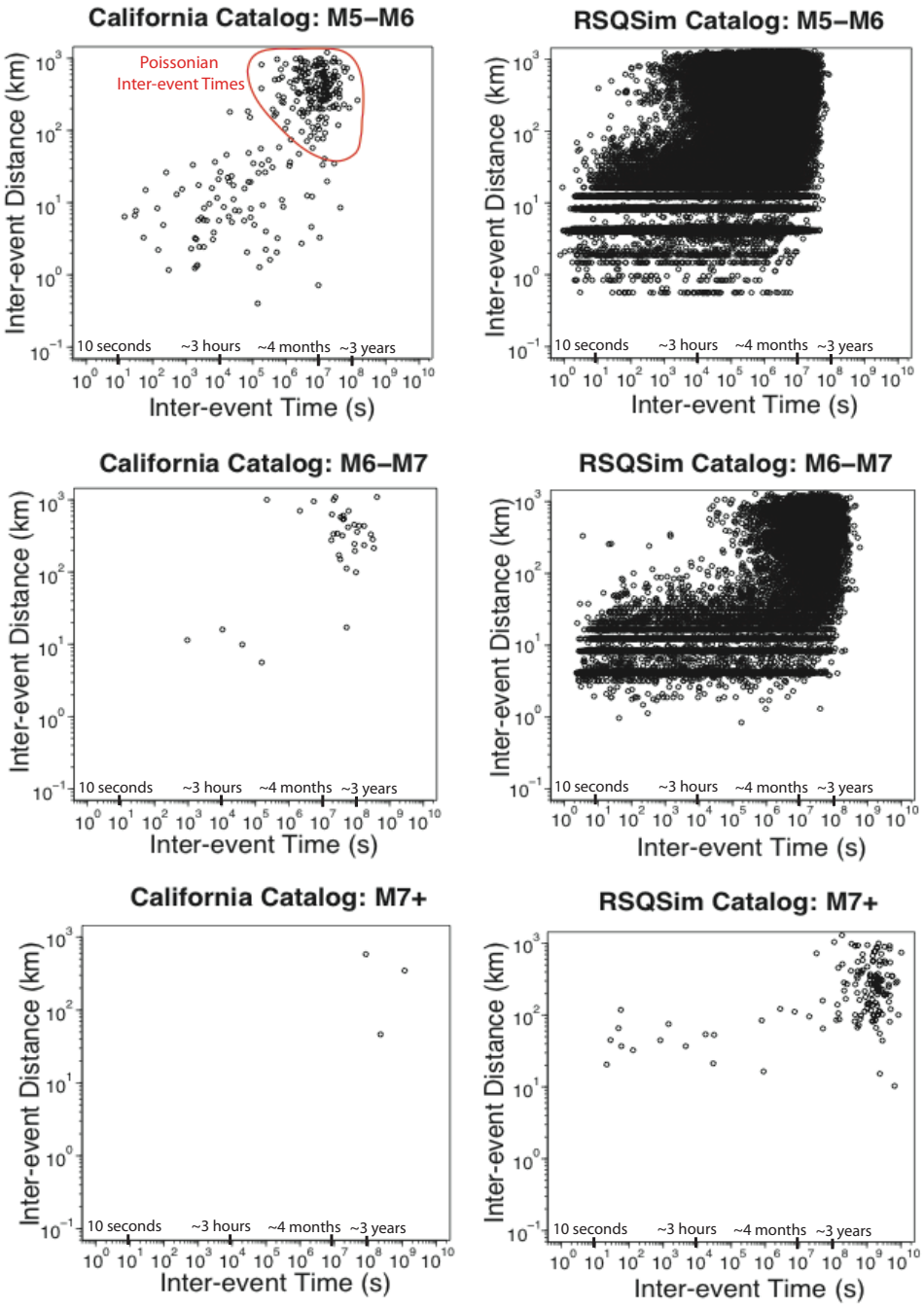

Figure 4.2: Inter-event distance probability distributions for the ANSS California catalog from 1911-2010 (left) and RSQSim UCERF2 catalog (right) for magnitude ranges M5-M6 (top), M6-M7 (middle) and M7+ (bottom). The inter-event times from the California M5-M6 catalog that fit an exponential (Poisson) distribution are circled in red to indicate the non-clustered events. The same pattern can be seen in the other catalogs as well. The striping of the RSQSim catalogs is due to the element size of the model. The distances between events are usually a multiple of the element size. Where there is enough data, the California catalogs show the same general trend as the UCERF2 catalog.

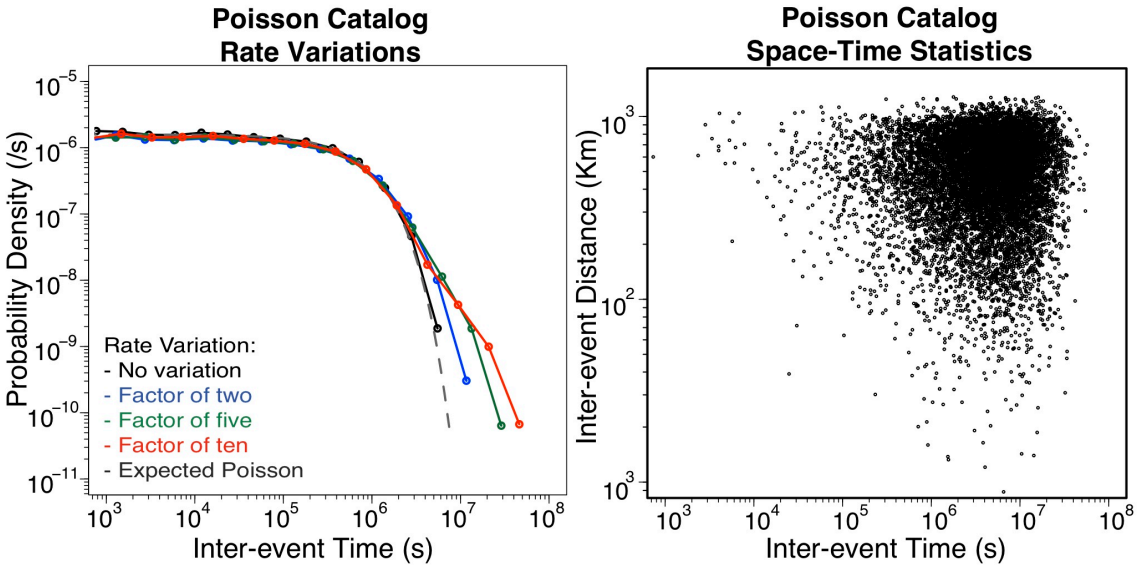

Figure 4.3: Inter-event time probability distribution (left) and inter-event distance distribution (right) for Poissonian catalogs. These catalogs were created by randomly drawing event times and locations, then combining catalogs with different event rates to simulate non-stationary Poisson event rates. In the left panel the dashed grey line gives the probability density for the stationary Poisson catalog. Increasing the amount of variation of the event rate within a catalog causes the 'tail' of the distribution (i.e. the end of the distribution at the longest inter-event times) to move outward, away from the stationary Poisson curve (grey dashed line). The colored lines (blue, green and red) are the inter-event time probability distributions of catalogs in order of increasing rate variation.

4.5 Clustering Probabilities of Large Events

For the primary purposes of this study, a large-event cluster has been defined simply as two or more earthquakes of $M \geq 7$ that occur within 4 years and one rupture length of apart. In simulations these large-event clusters usually consist of event pairs, but occasionally consist of three or more large events. The $M \geq 7$ threshold was chosen to study the events that present the greatest hazard, but statistics for $M \geq 6$ clusters are also presented in some sections. Several damaging earthquakes including the M6.9 Loma Prieta earthquake on October 17, 1989 and the M6.7 Northridge earthquake on January 17, 1994 fall below this M7 cut-off but caused a significant damage and loss of life. This

demonstrates the need to examine clustering over a wider magnitude range, as these slightly lower magnitude earthquakes still present a significant hazard.

A representative example of one such large-event cluster is shown in Figure 4.4. This example is from the UCERF3 'Paleoseismic' Catalog detailed in Chapter 3. In this example, the primary event in this large-event cluster is a *M*7.1 earthquake that nucleates on the 'Big Bend' section of the San Andreas Fault. A red star marks the nucleation point and the fault elements are colored by the total amount of slip in the event. Some elements are colored grey to indicate the locations of the aftershocks of the primary event. The secondary event in the cluster is a *M*7.4 earthquake that occurred on the Mojave Section of the San Andreas Fault just three years later. The rupture area of the primary clustered event is outlined in black to distinguish it from that of the secondary event, which re-ruptured several square kilometers of the rupture area of the primary event. The green elements to the left of the right-hand edge of the primary rupture zone mark the area that was re-ruptured. This area had less than one meter of slip before the secondary event occurred, which caused those elements to slip an additional two to three meters. Many of the secondary events in large-event clusters tend to rupture an adjacent section of the same fault that ruptured in the primary event. While the entire rupture of a primary event does not re-rupture in a secondary event, it is common for the secondary rupture to penetrate a few kilometers into the primary rupture area when secondary events occur along the same fault. Additionally, many of the secondary events in large-event clusters tend to rupture an adjacent section of the same fault that ruptured in the primary event, as illustrated in Figure 4.3.

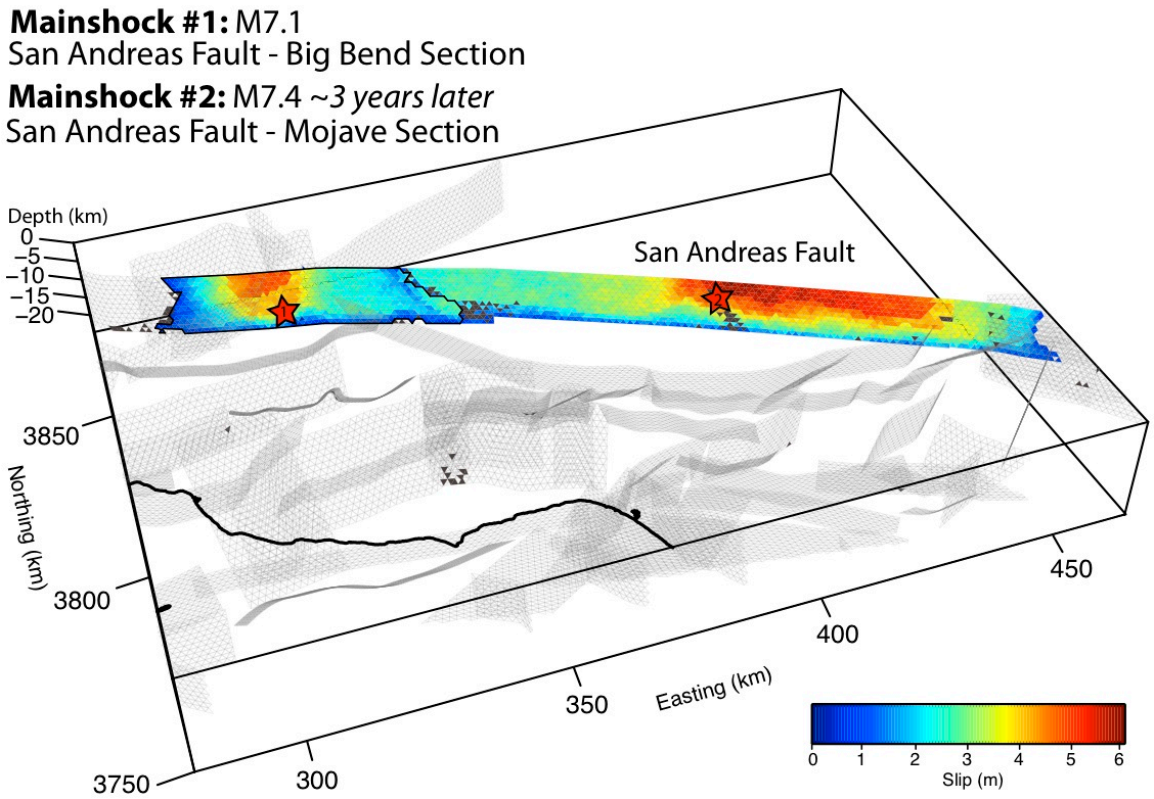

Figure 4.4: Large-event cluster example. Color indicates maximum slip on each element, dark grey elements indicate aftershock locations of Event #1, and red stars indicate hypocentral locations of each event. The rupture area of the first event in the cluster is outlined in black. The secondary event nucleates near the location of many of the aftershocks of the primary event.

Clustering statistics of the California catalogs, both simulated and instrumental, are parameterized in terms of the probability gain of large events ($M \geq M_{\min}$) following previous large events ($M \geq M_{\min}$) within some time period (Δt) due to the occurrence of large-event clusters. The probability gain ($G_{cluster}$) is defined as the ratio of the probability of a large-event cluster occurring ($P_{Observed}$) over the probability of random large events occurring ($P_{Poisson}$) within the same Δt. The equation for the probability gain is as follows:

$$G_{Cluster} = P_{Observed}/P_{Poisson} = \frac{N_{observed}}{N_{Total}} / \frac{N_{Poisson}}{N_{Total}}, \quad (4.2)$$

where, $N_{observed}$ is the total number of large events ($M \geq M_{min}$) after which at least one other large event ($M \geq M_{min}$) was observed in some time period (Δt) and within a radius of one rupture length (L) from the hypocenter of the primary $M \geq M_{min}$ event and $N_{Poisson}$ is the expected number of occurrences assuming a random (Poisson) distribution of events $M \geq M_{min}$ within the same primary event radius, and N_{Total} is the total number of $M \geq M_{min}$ events occurring within ΔT, the time span between the first and last event $M \geq M_{min}$ used for the analysis. $N_{observed}$ is determined by simply counting the number of observed intervals (Δt) that contain at least one additional event of $M \geq M_{min}$ with the distance L. The rupture lengths (L), and thus the search radius, ranged from 20 km to 400 km for the largest events but was 119 km on average. These probabilities are estimates that improve in accuracy as N_{Total} increases. For a stationary Poisson process, the expected occurrence of one or more additional events $M \geq M_{min}$ is given by the probability of one or more events in the interval Δt:

$$P = 1 - e^{-R_i \times \Delta t}, \quad (4.3)$$

where R_i is the long term rate of events $M \geq M_{min}$ within the specific radius (L) of each primary event i given by:

$$R_i = {}^{N_i}/_{\Delta T}, \quad (4.4)$$

where N_i is the number of $M \geq M_{min}$ events within the search radius. $N_{Poisson}$ is then simply the sum of the individual rates:

$$N_{Poisson} = \Sigma(1 - e^{-R_i \times \Delta t}). \quad (4.5)$$

For most calculations M_{min} = 7; however, statistics for $M \geq 6$ events are presented for comparison to the instrumental California catalog, which has only nine events above M7. In general, the probability gain indicates the increased probability of large events that occur in space-time clusters above what you would expect from a Poisson process. For example, $G_{cluster}$ of 1.50 for $M \geq 7$ events in 4 years, indicates that it is 150% more likely that another large event ($M \geq 7$) will occur within 4 years after a previous M≥7 event, compared to a Poisson process.

4.5.1 Effect of Constitutive Parameters on Clustering Probabilities

Several simulations were performed to test the effect of the rate-state friction parameter *a* on large-event clustering. In general, the rate of large events increases with increasing *a*-value. Clustering data for $M \geq 7$ events from four different simulated California catalogs with a-values ranging from 0.008 to 0.012 are shown in Table 4.1. All four of these catalogs have relatively high probability gains for large event clustering, indicating that there is a significant amount of clustering occurring. However, increasing the *a*-value by 50% decreases the probability of large-event clusters by almost 50%.

Model	Total # of events	Total #M_x of events	$P_{Observed}$ $M \geq 7$	$P_{Poisson}$ $M \geq 7$	$G_{Cluster}$ $M \geq 7$
a = 0.008	85,000	4,077	0.077	0.034	2.257
a = 0.009	120,000	4,551	0.086	0.042	2.031
a = 0.010	225,000	5,512	0.100	0.049	2.026
a = 0.012	1,300,000	6,314	0.094	0.053	1.767

Table 4.1: Clustering statistics of simulations from the UCERF2 fault model with varying a-values (the rate coefficient from the rate- and state- friction formulation) shown in Equation (1.1). $G_{cluster}$ is probability gain due to large-event clusters. M_x events in this case are M≥7 events. Increasing the a-value decreases the probability gain due to clustering of large events.

4.5.2 Clustering Probabilities for Different Magnitude Ranges

Dieterich and Richards-Dinger (2010) showed that RSQSim events display characteristic Omori (1894) aftershock decay, in which the rate of aftershock decay with time after a mainshock follows a power-law. Aftershocks from RSQSim catalogs typically have Omori slopes between p = -0.75 to p = -1.2 and are usually close to -0.9, depending on the fault model (Richards-Dinger and Dieterich, 2010). Additionally, the frequency distributions of the times between events in large-event clusters in RSQSim catalogs have similar power-law decay. Figure 4.5 shows the rate of $M \geq 7$ events with time after $M \geq 7$ mainshocks. The slope is -1.1, indicating that large-event clusters follow Omori's law with a decay rate that is similar to normal aftershock sequences.

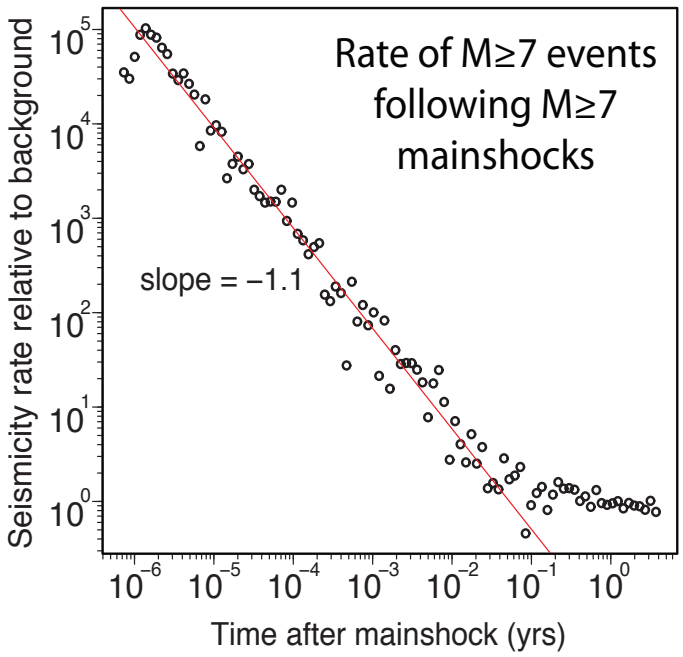

Figure 4.5: Seismicity rate decay for *M*7 events following *M*7 events, indicating that large-event clusters follow the Omori aftershock decay law.

The effect of event magnitude on the rate of large-event clustering seems to be significantly stronger than that of the rate-and state-dependent friction *a*-value. The clustering rate for $M \geq 6$ events is significantly higher than the rate for $M \geq 7$ events in the simulated catalogs. Clustering statistics for both UCERF catalogs and the ANSS California catalog from 1911-2015, for both magnitude ranges, are listed in Table 4.2. Overall, the probability gain for $M \geq 6$ events is nearly twice as high as that of $M \geq 7$ events. This is likely due to the fact that the occurrence rates of $M \geq 6$ events are significantly higher as well. The probability gains for $M \geq 7$ and $M \geq 6$ clustered events in the California catalog are surprisingly low compared to both simulated catalogs. However, the calculations for the instrumental catalog, particularly for the $M \geq 7$ events, were performed with so few events that these probabilities may not be valid. It was also necessary to make additional assumptions about the rupture lengths of the instrumental events in order to determine the clustering search radii. The rupture lengths (L), which were used as the search radius for the secondary events in large event clusters, were estimated from the magnitudes using the Wells and Coppersmith (1994) relationship for all slip types. Additionally, some of the $M \geq 6$ events in the California catalog may have occurred on faults that are not yet included in the UCERF fault models.

Catalog	# of M_x events	$P_{Observed}$	$P_{Poisson}$	$G_{Cluster}$
California M6+	**68**	**0.221**	**0.128**	**1.724**
RSQSim UCERF2 *M*6+	599,739	0.276	0.076	3.641
RSQSim UCERF3 *M*6+	*169,774*	0.194	0.068	2.850
California M7+	***9***	***0.111***	***0.095***	***1.171***
RSQSim UCERF2 M7+	*95,600*	*0.097*	*0.058*	*1.661*
RSQSim UCERF3 M7+	*49,935*	*0.074*	*0.053*	*1.416*

Table 4.2: Clustering statistics of the RSQSim UCERF3 catalog for clusters of magnitude M_x events, where x is either *M*≥7 or *M*≥6, compared to those of the ANSS California catalog from 1911-2015 years. Probability gains due to clustering of large events increase for smaller magnitude clusters. While the probability gains for the instrumental California catalog shows a similar trend for increases gains with magnitude, the calculation was performed with too few events to be considered significant.

4.5.3 Effect of Inter-Event Time on Clustering Probabilities

The inter-event time, the time between events for which they are considered clustered, is perhaps the most significant factor for clustering probabilities. Table 4.3 lists the clustering statistics for *M*≥7 and *M*≥6 events at 4 different inter-event times (Δt): 4 years, 1 year, 1 month, and 1 week. The probability gains due to shorter-term clustering increase dramatically for both magnitude ranges. Inter-event times of only 1 week reveal probability gains that are 54 and 110 times those of the longer-term clustering of 4 years for the *M*≥7 and *M*≥6 ranges respectively, indicating very strong short-term clustering probabilities of large events compared with random chance. This short-term clustering is primarily do to triggering by stress transfer and the time delay of nucleation inherent in rate- and state-dependent friction. This effect is especially pronounced when compared to clustered pairs with longer (up to 4 year) inter-event times because, while the shorter inter-event time clusters get counted with the 4-year clusters, the Poissonian probability is significantly decreased for the short inter-event times.

Catalog	Δt	# of M_x events	$P_{Observed}$	$P_{Poisson}$	$G_{Cluster}$
UCERF3 M6+	4 years	169,774	0.194	0.068	2.850
UCERF3 M6+	1 year	169,774	0.156	0.019	8.266
UCERF3 M6+	1 month	169,774	0.128	0.002	78.419
UCERF3 M6+	1 week	169,774	0.117	0.0004	312.222
UCERF3 M7+	4 years	*49,935*	*0.074*	*0.053*	*1.416*
UCERF3 M7+	1 year	49,935	0.040	0.014	2.947
UCERF3 M7+	1 month	49,935	0.024	0.001	20.435
UCERF3 M7+	1 week	49,935	0.020	0.0003	75.986

Table 4.3: Clustering statistics of the UCERF3 catalog for M_x = M≥7 or M≥6, at 4 different inter-event times (Δt), the time between events in large-event clusters, of 4 years, 1 year, 1 month, and 1 week. Probability gains increase dramatically when the inter-event time between clustered events is decreased, indicating very strong short-term clustering and potential triggering.

4.5.4 Effect of Fault Geometry on Clustering Probabilities

Fault geometry also acts as a control on clustering rates. Dieterich & Richards-Dinger (2010) showed that varying the small-scale complexity of fault models produces different clustering probabilities. They found that models of smooth faults, faults with fractal roughness, and faults with fractal segmentation all have different clustering probabilities and that the amplitude of segmentation was the strongest influence on the rate of large-event clusters. The data from Dieterich and Richards-Dinger (2010) has been modified to show the clustering probabilities and the probability gain due to that clustering, and the results are listed in Table 4.4. The amplitude of the roughness (β) of the fractally-rough and fractally-segmented models is the RMS slope of the deviations from a planar surface. The UCERF2 and UCERF3 fault models have much more complicated geometry on a large scale than the models presented in Dieterich and Richards Dinger (2010), including curved and dipping faults, and fault zones with bends and step-overs. However, the clustering probability data from those cannot be directly

compared to this study because the simulations have too many differences, including fault slip rates and values of the rate- and state-dependent friction constitutive parameters. The models with more complex fault geometry display the highest clustering probabilities. Fault segmentation, either fractal or in the form of fault branches and stepovers, acts as a barrier to thoroughgoing ruptures which results in delayed rupture propagation. This delayed rupture propagation is the cause of many large event clusters, as two events occur in place of one continuous event.

Model	**Total #M_x of events**	**$P_{Observed}$ M≥7**	**$P_{Poisson}$ M≥7**	**$G_{Cluster}$ M≥7**
Smooth Faults	196	0.177	0.075	2.360
Fractal Roughness (β=0.10)	237	0.190	0.090	2.108
Fractal Segmentation (β=0.02)	221	0.180	0.084	2.133
Fractal Segmentation (β=0.04)	274	0.395	0.103	3.814

Table 4.4: Clustering statistics for Mx = M≥7 event from the RSQSim California fault model compared to those of several different models of idealized fault systems. Table modified from Dieterich and Richards-Dinger (2010) to include data from the California fault models. Models with greater fault complexity tend to exhibit greater probability gains from large-event clustering.

Within the California fault models there is also a spatial variation to the clustering of large events. The large-event clusters often develop spontaneously on continuous fault segments, but clustering rates are higher in areas of structural complexity. Large-event clustering increases with separation distance between discontinuous elements as well as geometric complexity. Figures 4.6 and 4.7 illustrate the variation in probability gain due to large-event clustering in space for M≥7 and M≥6 events respectively. It is colored by the ratio of the probability of clustered large events (counted in 100x100 km boxes) to the probability of expected large events from a Poisson distribution with the same rate. Both

figures have the same color scale for comparison. The dark blue colors are the areas where there were fewer than expected large-event clusters (for a Poisson process) and the green to red boxes are the areas with higher than expected numbers of clustered events. The dark blue areas (where the probability gain is less than one) are areas with events that are more periodic, rather than clustered. The bright red boxes have nearly one hundred times as many clustered events than expected from the Poisson model of occurrence. The overall probability gains are higher for $M \geq 6$ events primarily because the rate of clustering is higher for those smaller magnitude events. This higher rate of triggering may be due, in part, to the difference in aspect ratio between $M \geq 7$ and $7 < M \geq 6$ events. This results in greater stress transfer, averaged over the potential areas of secondary $7 < M \geq 6$ events, than for secondary $M \geq 7$ events. The width to length ratio for the $7 < M \geq 6$ events in RSQSim catalogs is close to one and these nearly square ruptures cover a much smaller area of the fault than the long, $M \geq 7$ events. Since secondary clustered events tend to occur off the ends of the primary event ruptures, the potential clustering area is smaller for larger events.

Additionally, lower slip rate faults have larger clustering probability gains than higher slip rate faults. The variation in clustering probability gain for faults with different slip rates is plotted in Figure 4.8 for the UCERF3 catalog, for $M \geq 6$ (blue) and $M \geq 7$ (red) events. The probability gains are sorted into 5mm/year bins and the solid lines indicate the total probability gains for both magnitude ranges. Not only are the probability gains higher for $M \geq 6$ events than the $M \geq 7$ events for each slip rate bin, but the variation in and range of probability gains is much higher as well (0.4-5.1 for $M \geq 6$ events and 0.6 to 2.5

for $M \geq 7$ events). The total probability gain for $M \geq 6$ events however seems to be dominated by the high probability gain (5.1) at slip rates less than 5mm/year. This dependence on fault slip rate is expected because higher slip rates will tend to have proportionally higher earthquake rates, which increases the Poisson probability for random event clusters and decreases the probability gain.

Overall, the probability gains appear to be higher for all magnitudes in the areas where faults have lower slip rates and where faults have significant bends, branches, and stepovers. This indicates that there is more clustering in areas of greater fault complexity, where there are significant barriers to through-going ruptures. These geometric barriers tend to increase clustering because ruptures are forced to stop and re-nucleate on another section of the fault, producing two (or more) earthquakes instead of one.

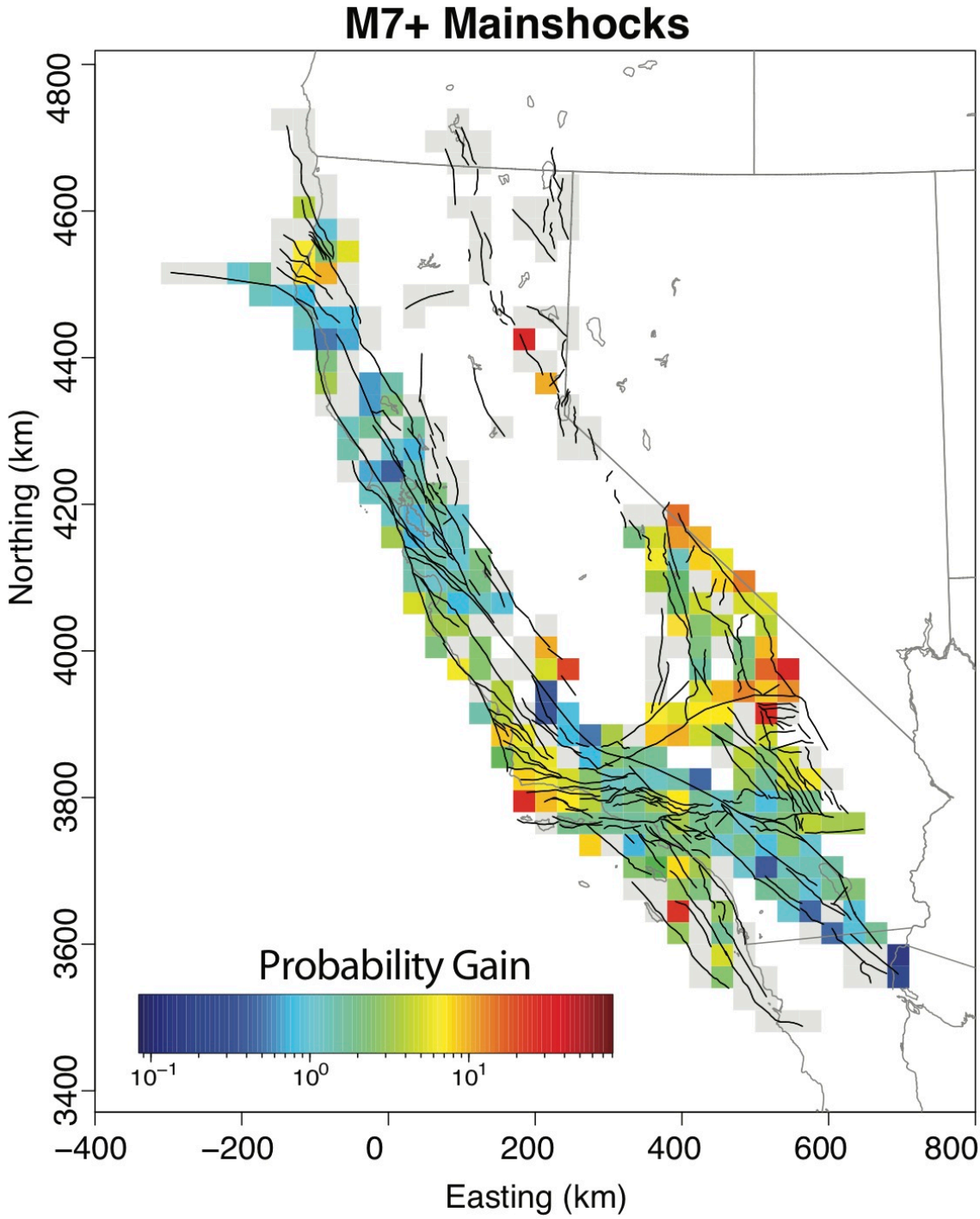

Figure 4.6: Spatial variation in clustering rates of M≥7 events, indicated by the ratio of clustered events to Poissonian events in each 30km by 30km box. Clustered events on these maps are defined as being within 4 years time and 30 km hypocentral distance. Clustering of M≥7 events tends to occur in areas of greater fault complexity such as branches or step-overs, however large-event clusters also occur on long, simpler fault sections.

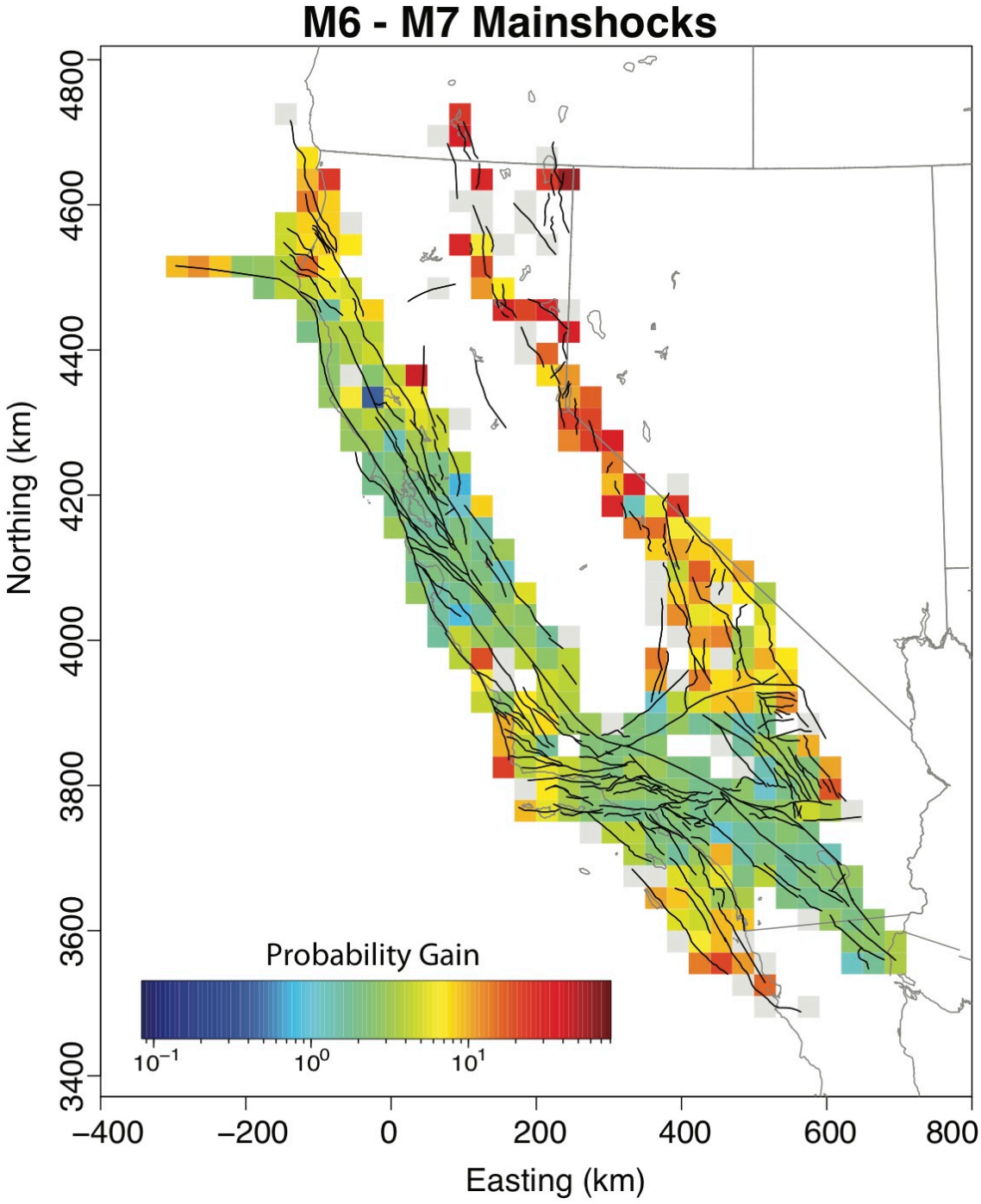

Figure 4.7: Spatial variation in clustering rates of M6 to M7 events, indicated by the ratio of clustered events to Poissonian events in each 30km by 30km box. Clustered events on these maps are defined as being within 4 years time and 30 km hypocentral distance. Clustering of M6 to M7 events tends to occur much more often than for M≥7 events, but is still greatest in areas of greater fault complexity such as branches or step-overs, however large-event clusters also occur on long, simpler fault sections.

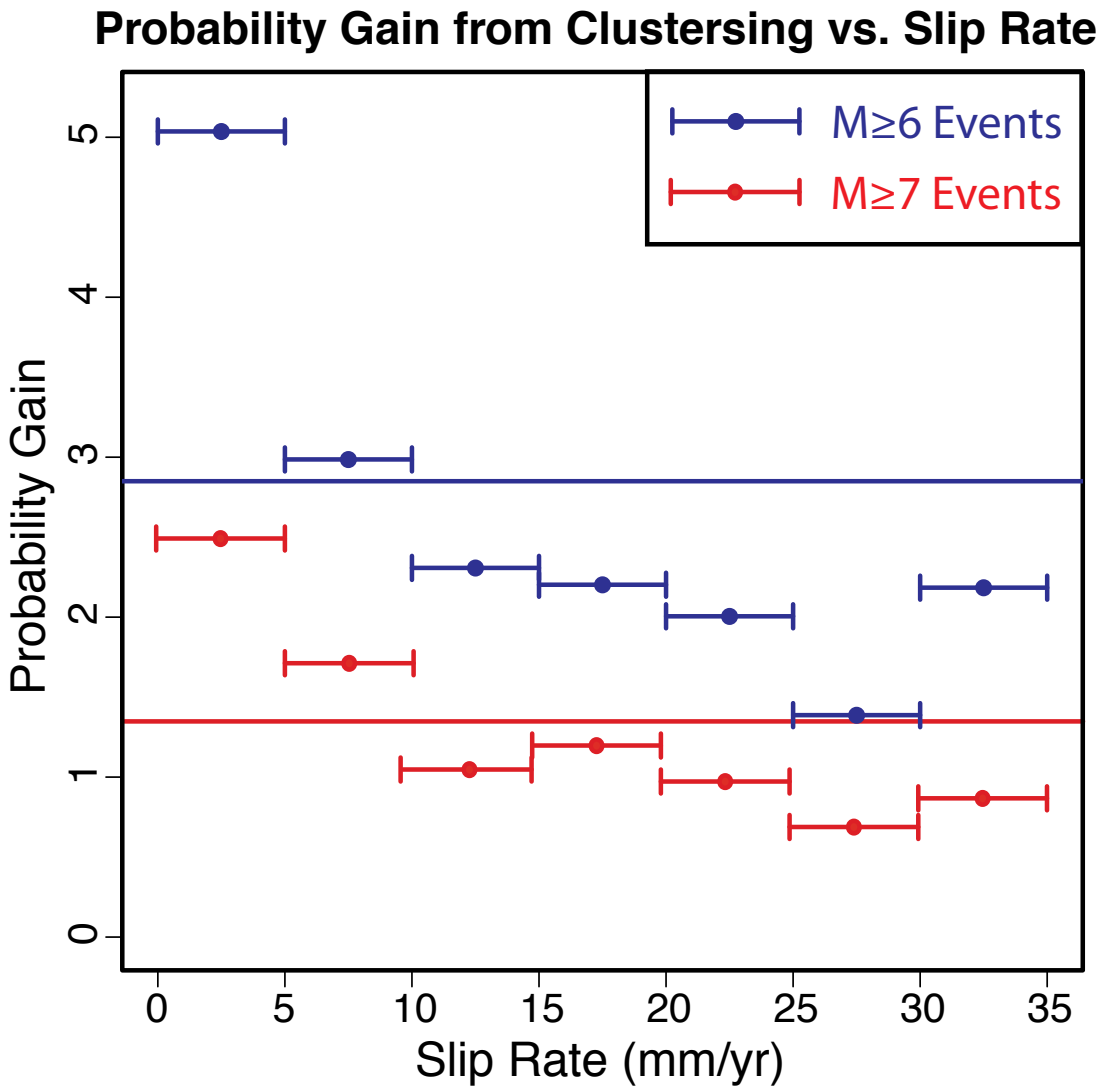

Figure 4.8: Probability gain due to clustering of large events (M≥7 in red and M≥6 in blue) for faults with different slip rates. The probability gains are sorted in 5mm/year bins as indicated by the horizontal bars, and the solid lines indicate the total probability gains for both magnitude ranges. The probability gains are higher (for both magnitude ranges) for clustered events on faults with lower slip rates.

4.6 Conclusions

Instrumental earthquake catalogs are much too short to directly characterize clustering statistics of potentially damaging earthquakes in specific regions like Southern California. However, earthquake simulators show a great deal of promise for this objective. Simulated catalogs have aftershocks with realistic Omori trends and the inter-event time and space-time statistics of smaller earthquakes in the RSQSim catalogs are in good agreement with the California catalogs. We therefore conclude that the RSQSim catalogs can provide an appropriate basis for investigation of large-event clustering. Probabilities of large-event clusters observed in RSQSim catalogs are significantly higher than the probabilities for a truly random process, establishing that many secondary events in large-event clustered are not only related to, but are actually triggered by the primary events.

Clustering probabilities vary depending on several factors including fault complexity, rate- and state-dependent friction constitutive parameters, slip rate, and clustered event magnitude. Geometrically complex fault systems show more clustering behavior than simple systems. Increasing the value of the rate variable a in the rate- and state-friction equation yields higher probability gains due to clustering. Finally, clustering probabilities are much higher for smaller events ($M6$ vs. $M7$). Overall, these results could provide valuable information related to forecasting large events. Large-event clusters like the one shown in Figure 4.4 would be devastating for Southern California, especially for highly populated cities like Los Angeles, which would likely experience strong ground shaking from both events.

4.7 References

Ando, M. (1975). Source mechanisms and tectonic significance of historical earthquakes along the Nankai Trough, Japan. *Tectonophysics*, *27*(2), 119-140.

Davidsen, J., & Goltz, C. (2004). Are seismic waiting time distributions universal?. *Geophysical Research Letters*, *31*(21).

Dieterich, JH. (1994), A constitutive law for rate of earthquake production and its application to earthquake clustering, J. Geophysics. Res., 99(B2), 2601–2618, doi:10.1029/93JB02581.

Dieterich, J.H., and Kilgore, B.D. (1996). Imaging surface contacts: power law contact distributions and contact stresses in quartz, calcite, glass, and acrylic plastic, Tectonophysics, 256, 219-239.

Dieterich, J. H., and Richards-Dinger, K. B. (2010), Earthquake recurrence in simulated fault systems, Pure Appl. Geophysics., 167,1087–1104, doi:10.1007/s00024-010-0094-0.

Field, E. H., Arrowsmith, R. J., Biasi, G. P., Bird, P., Dawson, T. E., Felzer, K. R., ... & Zeng, Y. (2014). Uniform California Earthquake Rupture Forecast, Version 3 (UCERF3)—The Time-Independent Model. *Bulletin of the Seismological Society of America*, *104*(3), 1122-1180.

Hainzl, S., Zöller, G. and Scherbaum, F. (2003). Earthquake clusters resulting from delayed rupture propagation in finite fault segments, J. Geophysics. Res., 108(B1), 2013, doi:10.1029/2001JB000610.

Helmstetter, A., Kagan, Y. Y., & Jackson, D. D. (2005). Importance of small earthquakes for stress transfers and earthquake triggering. *Journal of Geophysical Research: Solid Earth (1978–2012)*, *110*(B5).

Kagan, Y.Y. and Jackson, D.D. (1991). Long-term earthquake clustering, Geophysics. J. Int., 104(1), 117–133.

Kagan, Y.Y. and Jackson, D.D. (1999). Worldwide doublets of large shallow earthquakes, Bull. Seism. Soc. Am., 89(5), 1147–1155.

Kagan, Y.Y. and Knopoff, L. (1978). Statistical study of the occurrence of shallow earthquakes, Geophysics. J. R. Astron. Soc., 55, 67-86.

King, G. C. P., Stein, R. S. and Lin, J. (1994). Static stress changes and the triggering of earthquakes, Bull. Seismol. Soc. Amer., 84, 935-953.

NCEDC (2014), Northern California Earthquake Data Center. UC Berkeley Seismological Laboratory. Dataset. doi:10.7932/NCEDC.

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association*,*83*(401), 9-27.

Ogata, Y., and Abe, K. (1991). Some statistical features of the long-term variation of the global and regional seismic activity. Intn. Statistical Review, 59, 139-162.

Omori, F. (1894). Investigation of Aftershocks, Rep. Earthq. Inv. Comm., 2, 103–39.

Reasenberg, P. (1985). Second-order moment of central California seismicity, 1969–1982, J. Geophysics. Res., 90, 5479–5495.

Reasenberg, P. A., & Jones, L. M. (1989). Earthquake hazard after a mainshock in California. *Science*, *243*(4895), 1173-1176.

Richards-Dinger, K., and J. H. Dieterich (2012), RSQSim Earthquake Simulator, Seismological Research Letters, v. 83, no. 6, p. 983-990, 2012.

Scharer, K.M., Biasi, G.P., Weldon, R.J. (2011). A reevaluation of the Pallett Creek earthquake chronology based on new AMS radiocarbon dates, San Andreas Fault, California, Journal of Geophysical Research, 116, B12111.

Stein, R. S., Barka, A. A. and Dieterich, J. H. (1997). Progressive failure on the North Anatolian fault since 1939 by earthquake stress triggering, Geophysics Journal International, 128, 594–604.

Toda, S., and Stein, R. S. (2002). Response of the San Andreas Fault to the 1983 Coalinga-Nuñez earthquakes: An application of interaction-based probabilities for Parkfield, J. Geophysics. Res., 107(B6), doi:10.1029/2001JB000172.

Utsu, T. (1961). A statistical study on the occurrence of aftershocks. *Geophysical Magazine*, *30*(4).

Utsu, T., & Ogata, Y. (1995). The centenary of the Omori formula for a decay law of aftershock activity. *Journal of Physics of the Earth*, *43*(1), 1-33.

Wang, Q., Jackson, D. D. and Zhuang, J. (2010), Missing links in earthquake clustering models, Geophysics. Res. Lett., 37, L21307, doi:10.1029/2010GL044858.

Wells, D. L., & Coppersmith, K. J. (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bulletin of the Seismological Society of America*, *84*(4), 974-1002.

Chapter 5

Earthquake Probabilities from Aftershock Sequences in Simulated Catalogs

5.1 Abstract

The occurrence of large-event clusters, in which a second event of similar magnitude follows one large event within a few years or less, indicates that the probability of additional large events is temporarily increased after a large event occurs. While understanding of this increased probability lends valuable information about changes in earthquake hazard, stress changes and aftershock rates may give additional predictive information. The earthquake simulator RSQSim generates simulated catalogs that include foreshocks, aftershocks, and occasional large-event clusters. In the simulations, we observe several potential indicators of impending large, secondary events comparable in magnitude to the prior mainshocks in the simulated catalogs. When a large-event cluster occurs, the primary event has an especially productive aftershock sequence, which on average is roughly double the aftershock rates for non-clustered mainshocks. Additionally, the aftershock locations of the first event in a cluster appear to correlate with the hypocentral location of the next large event in the cluster. We find that aftershock rate is a proxy for the stress state of the faults. Also, in large event clusters aftershock sequences transition into foreshock sequences, wherein the average event rate

increases prior to the impending large event in the cluster. These increased event rates prior to the second event in a cluster follow an inverse Omori's law, which is characteristic of foreshock sequences in nature. Finally, the aftershock/foreshock locations migrate toward the point of nucleation of the next large event in the cluster. Clustering probabilities based on aftershock rates have better predictive power than those obtained from Omori aftershock and Gutenberg-Richter magnitude frequency laws, which suggests that the high aftershock rates indicate near-critical stresses for failure in a large earthquake.

5.2 Introduction

Various possible earthquake precursors have been studied including seismic velocity changes, earthquake swarms or micro-earthquakes, uplift, preslip, tilt-strain anomalies, and ground water changes (Geller, 1997). Triggering and prediction studies have attempted to isolate precursors and quantify earthquake probabilities by studying seismicity, ground deformation, and Coulomb stress changes (Freed, 2005). No reliable prediction schemes have been developed however, because patterns of possible precursors vary so greatly between events (Geller, 1997). One problem in the search for possible precursors is the lack of data, especially for very large events. Earthquake simulations in RSQSim may be able to fill this gap and provide valuable statistics on earthquake occurrence.

Several studies have incorporated stress transfer calculations into seismic hazard assessments (Dieterich, 1988; King et al., 1994; Stein et al., 1997; Hardebeck et al., 1998;

Parsons et al., 1999). Some studies have explored cases of progressive earthquake failure in which successive events rupture a nearby section of the same fault (Richter, 1958; Dewey, 1976; Jackson & McKenzie, 1988; Scholz, 1990). Calculations based on Coulomb failure stress changes due to earthquakes on the North Anatolian fault were used by Stein et al. (1997) to estimate long-term earthquake probabilities in Turkey. Other studies have used Coulomb stress changes to determine whether certain faults were brought closer to failure by slip on nearby faults. King et al. (1994) showed that the *M*7.4 Landers earthquake raised the stress at the location of the *M*6.5 Big Bear earthquake which occurred just three hours later, and that both of these events raised the stress on a section of the southern San Andreas Fault enough to advance the tectonic stressing clock on the next event by a decade. While Coulomb stress transfer provides valuable insights for understanding the factors that control the clustering of large earthquakes, there are still only a few examples of Coulomb stress changes being used to forecast the occurrence of future large earthquakes. Furthermore, Coulomb stress calculations do not explain the statistics of clustering, nor their temporal characteristics (such as Omori-type time dependence).

5.3 Aftershock Sequences of Large-Event clusters

5.3.1 Highly Productive Aftershock Sequences

As demonstrated in Chapter 4, RSQSim simulations produce clusters of large events in California at rates consistent with naturally occurring clusters in California. In simulations with a highly idealized fault system, Dieterich and Richards-Dinger (2010)

report that large event clusters appear to correlate with highly productive aftershock sequences following the first event in the cluster. This chapter examines that possibility in greater detail by compiling aftershock statistics for clustered and non-clustered events from long simulations with California fault models. Figure 5.1 illustrates the slip distributions and aftershock locations of a large-event cluster consisting of two earthquakes M7.2 and M7.6 on a section of the San Andreas Fault and, for comparison, an independent M7.2 earthquake on the same fault section. This particular independent event was chosen for comparison because the slip distribution, i.e. the amount of slip on each individual element, was 99% similar to the primary event in the large-event cluster. This similarity was determined by comparing the normalized dot products of the vector of slip at each element for each event. The primary event in the cluster, an *M*7.2 on the Southern San Andreas Fault, was followed by 32 $M \geq 4.9$ aftershocks in the 29.9 days prior the occurrence of the *M*7.6 secondary event. Not only was this aftershock sequence highly productive (with over 30 large aftershocks in less than a month) but also 97% of the aftershocks occurred to the south of the rupture area of the primary event near the location of the nucleation point of the secondary event (indicated by a red star on Figure 5.1). A similar sequence was observed for the Elmore Ranch and Superstition Hills earthquake sequence in 1987, in which the *M*6.2 Elmore Ranch earthquake (November 23rd) triggered the *M*6.6 Superstition Hills earthquake (November 24th) on an intersecting fault. The Elmore Ranch earthquake had a particularly energetic aftershock sequence and the aftershocks were primarily located near the intersection of the faults, which was also the nucleation location of the Superstition Hills earthquake 11.4 hours later (Hudnut et

al., 1989). Conversely, the independent RSQSim event only had three aftershocks in the same time period. In general, the aftershocks of independent events are relatively evenly distributed around the ends of the ruptures. While Figure 5.1 is a rather extreme example of the difference in aftershock productivity for a clustered and an independent event, a compilation of aftershock data shows that clustered events have, on average, twice as many aftershocks as independent events. Figure 5.2 illustrates the aftershock decay for all $M \geq 7$ independent events (blue) and $M \geq 7$ primary clustered events (red) for the half-million year UCERF3 catalog. Aftershock data for the entire state are stacked in time bins based on the amount of time between each aftershock and the mainshock for over 9000 large-event clusters and over 80,000 independent $M \geq 7$ events. Not only do the clustered events produce more aftershocks, but also the decay rate of those aftershocks is also somewhat lower than that of the aftershocks of independent events. The difference in aftershock productivity between clustered and independent events is apparent within just a few minutes after a mainshock. Figure 5.3 shows the average cumulative number of aftershocks per mainshock with time after the mainshock for both the primary clustered events (red) and independent $M \geq 7$ events (blue), from the same set of stacked events in Figure 5.2. Aftershocks are only counted within one rupture length from the ends of the rupture area for each mainshock. The number of aftershocks between the two distributions of events diverges within 2-3 minutes after mainshock occurrence. In general, events that become part of a large-event cluster have higher aftershock rates than independent events. Results show that simulated California events of $M \geq 7$ that are followed by another $M \geq 7$ event within 4 years have more productive aftershock

sequences than other $M \geq 7$ events that occur independently. The differences in productivity suggest that aftershock sequences of large events may provide useful information about the likelihood of additional large events occurring within a few years, as well as the general locations of those additional events.

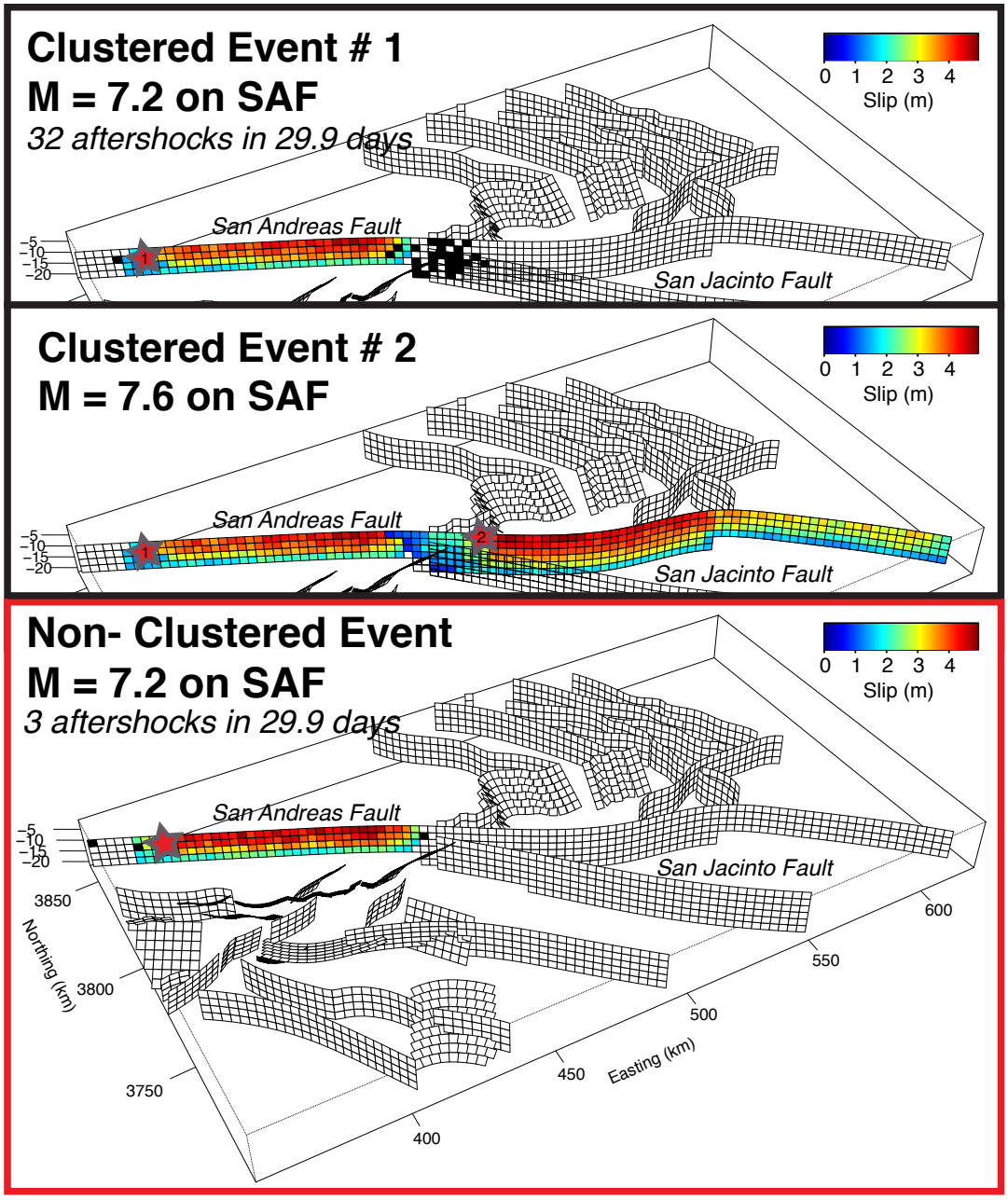

Figure 5.1: Large-event cluster on the Southern San Andreas Fault (top – black outline) and non-clustered or independent event - not followed by another large event within 4 years (bottom – red outline) that is nearly identical (99% similar slip on individual elements) to the first event in the large-event cluster. Colors represent maximum slip on each fault element and black elements indicate the locations of aftershocks. Not only does the clustered event have significantly more aftershocks, but also those aftershocks are located near the hypocenter of the secondary event in the cluster.

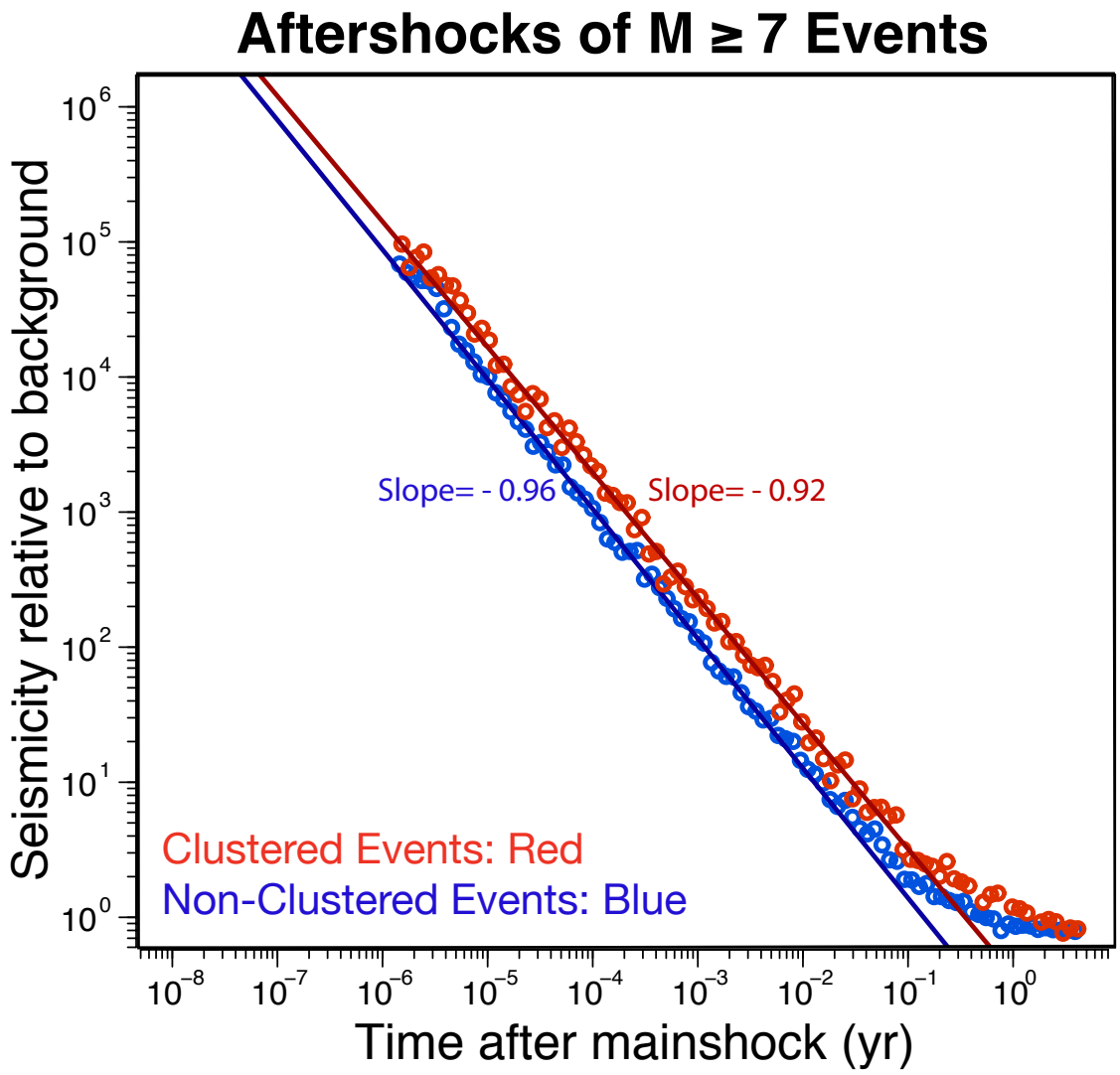

Figure 5.2: Aftershock productivity of Clustered Events (red) and Non-Clustered Events (blue) with time after the mainshock. Includes stacked data, binned by the amount of time between each aftershock and the mainshock, for over 9000 large-event clusters and over 80,000 independent M≥7 events. The aftershock rate is higher for clustered events and the decay rate is lower. This difference in productivity with time could provide additional information about the short-term probabilities of large events.

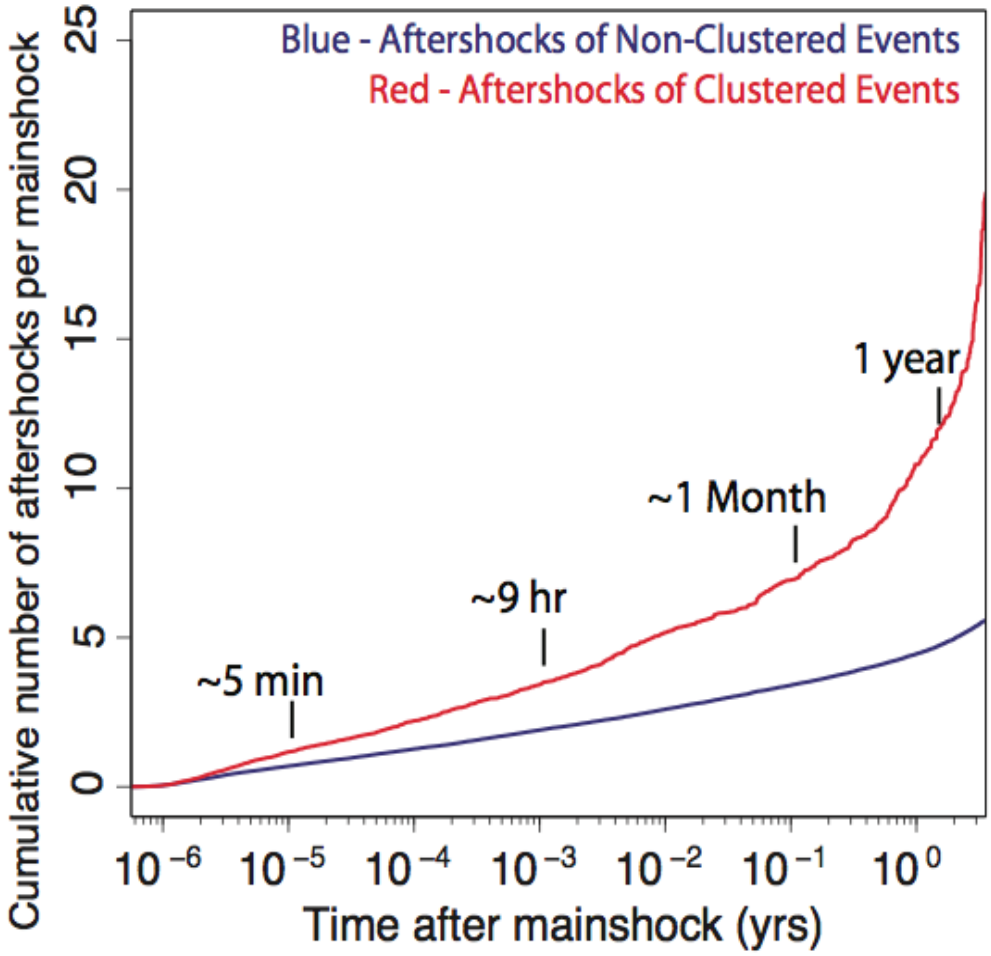

Figure 5.3: Cumulative number of aftershocks with time from mainshock. The blue line represents the aftershocks of non-clustered events and the red line represents the aftershocks of clustered events. Includes stacked data for over 9000 large-event clusters and over 80,000 independent M≥7 events. The difference in aftershock productivity is apparent within a few minutes after the mainshock, suggesting that increases in short-term earthquake probability could be recognized based on real-time aftershock data.

5.3.2 Aftershock Migration and Foreshock Evolution

In addition to elevated aftershock rates following the primary events of large-event clusters, the aftershocks also transition to foreshocks with accelerating rates prior to the impending secondary large event of the cluster. In other words, following a normal

aftershock sequence in which the rate of smaller events increases dramatically after a mainshock and decays over time to the background seismicity level, following Omori's law (Equation 4.4), the average event rate (of aftershocks of the primary clustered events) increases prior to the secondary event in the large-event cluster. Figure 5.4 shows the cumulative number of aftershocks (counted only within one rupture length from the rupture ends of the mainshock) per mainshock with time until the secondary event, for clustered events that had 3 to 4 years between them in the half-a-million year UCERF3 catalog. For this plot, in order to stack the aftershock data during the variable 3 to 4 year intervals, the time between the primary and secondary events is normalized by the inter-event time of each clustered pair for which aftershocks were counted. The aftershock rates of the primary events in large-event clusters show typical Omori decay before the seismicity rate accelerates**.** These increased event rates follow an inverse Omori's law, which is characteristic of foreshocks.

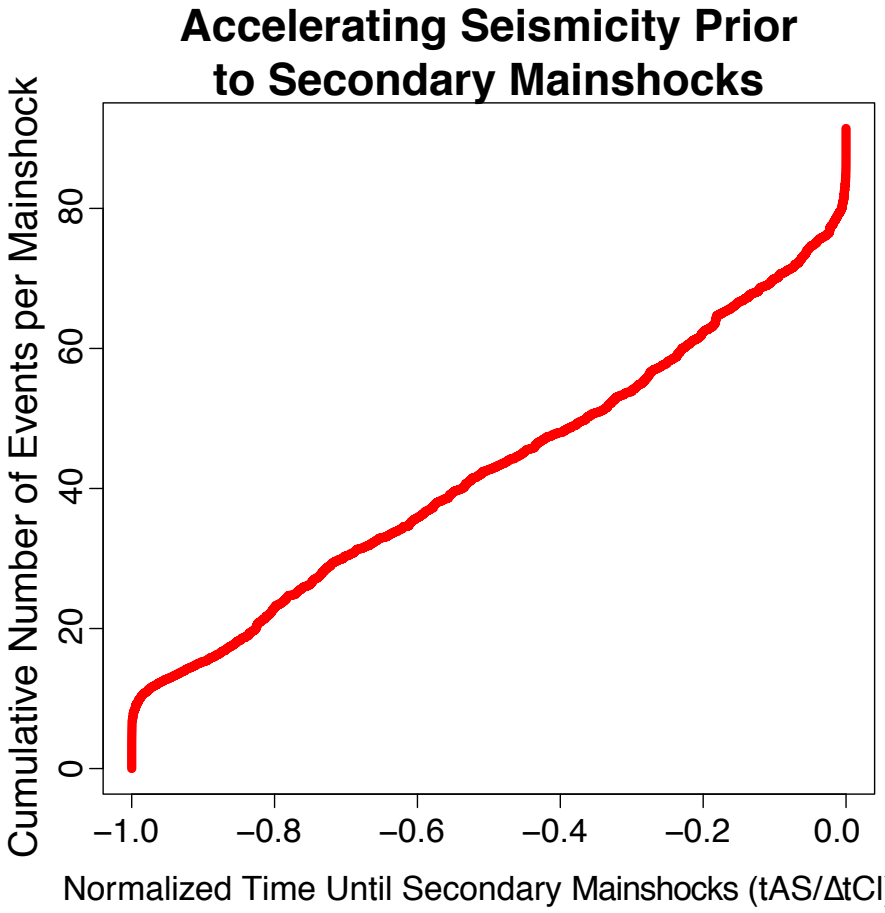

Figure 5.4: Cumulative aftershocks per mainshock (the primary event in each large-event cluster) with time leading up to the time of the secondary event in each large-event cluster. Time is normalized by the amount of time between the primary and secondary events in each large-event cluster. Data has been stacked for large-event clusters with 3-4 years between them. Aftershock rates show typical Omori decay (like Figure 4.3) and start to level off before increasing in an inverse Omori trend, characteristic of foreshocks, shortly before the secondary event occurs.

Furthermore, as the seismicity transitions from an aftershock to a foreshock sequence, it localizes near the point of nucleation of the secondary event in a large-event cluster. As illustrated in the example in Figure 5.1, the aftershocks of the primary event in a large-event cluster tend to correlate with the location of the next large event in the cluster. Not only do most secondary clustered events nucleate at the outer edge of the

aftershock zone of the primary clustered events, but the aftershocks also migrate in space towards the secondary nucleation location with time. Figure 5.5 shows the mean distance between the aftershocks of primary clustered events and the nucleation location of the secondary clustered event with time leading up to the nucleation time of the secondary clustered event. Once again, this was calculated for stacked clustered events that had 3 to 4 years between them, and the time is normalized by the specific inter-event time for each clustered pair. Aftershocks are counted within 50 km of the end of the primary rupture that is nearest to the nucleation point of the secondary event. On average the aftershocks that occur closer in time to the second event are also closer in space and these aftershocks clearly localize toward the point of nucleation of the next event in a large-event cluster.

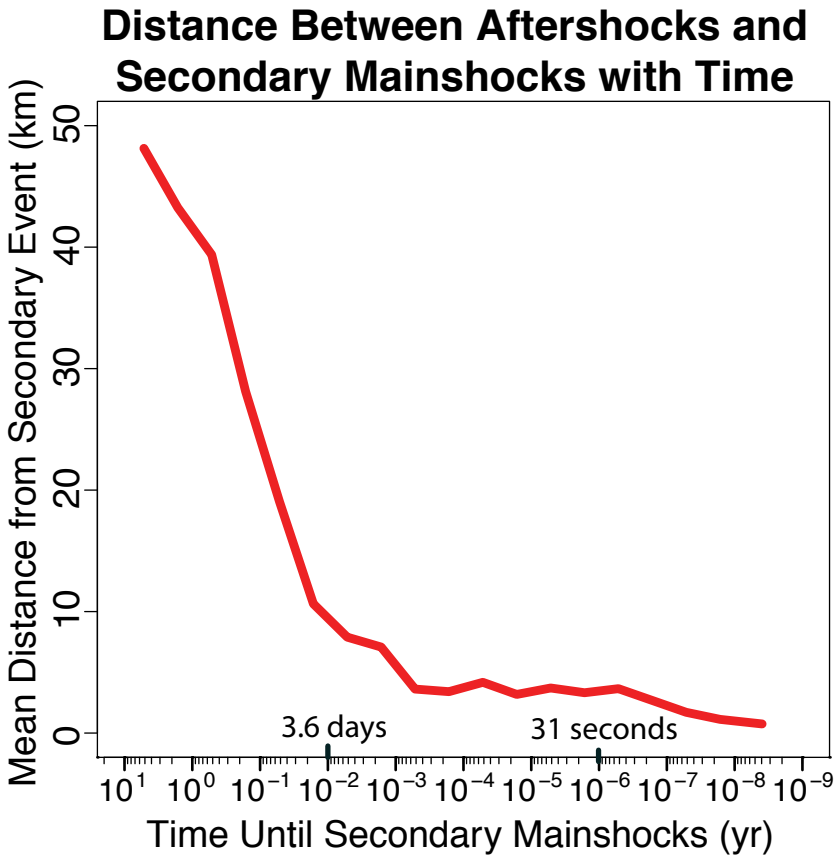

Figure 5.5: Distance between the aftershocks of the primary events in large-event clusters and the location of the secondary event in the cluster with time leading up to the time of the secondary event. Aftershocks are counted within 50 km of the end of the primary rupture that is nearest to the nucleation point of the secondary event. As the aftershocks transition into foreshocks approaching the time of occurrence of the secondary event, the foreshocks occur closer in space to the hypocentral location of that event as well.

5.3.3 Effect of Initial Stresses on Clustering

As illustrated by the large-event cluster in Figure 5.1, it is fairly common for large-event clusters to progressively rupture the next section of the fault. The existence of these large-event clusters, as well as the progressive rupture of adjacent sections, suggests that stress transfer from the primary event triggers the secondary event in the cluster. As previously mentioned, Stein et al. (1997) report 3-month to 30-year periods between stress transfer and progressive triggering of events along the North Anatolian Fault,

indicating not only that the time periods between clustered events nature are roughly consistent with those of simulated catalogs, but that static stress transfer may be a reasonable mechanism for the triggering of secondary events in large-event clusters. Figure 5.6 illustrates the change in the inter-event time between clustered events based on the magnitude of the Coulomb stress change on the rupture area of the secondary clustered event due to the occurrence of the primary clustered event. The Coulomb stress change reflects changes in both shear stress and normal stress and takes into account fault geometry, as well as the sense of slip. This absolute Coulomb stress was calculated based on Equation 1 from King et al. (1994), but we use a modified form of the equation that ignores pore fluid pressures, which is as follows:

$$S = \tau - \mu_0 * \sigma, \tag{5.1}$$

where S is the Coulomb stress on the fault or fault element in question, τ is the shear stress, μ_0 is the steady state coefficient of friction (0.6 for all simulations discussed in this study), and σ is the normal stress. The change in Coulomb stress is calculated as follows:

$$\Delta S = \Delta\tau - \mu_0 * \Delta\sigma, \tag{5.2}$$

Where $\Delta\tau$ is the change in shear stress, and $\Delta\sigma$ is the change in normal stress. This is calculated for each fault element, then averaged over the entire rupture area of the secondary event. The inter-event times of clustered pairs correspond with the magnitude of the stress perturbations from primary clustered events. On average, secondary events that experience greater Coulomb stress increases occur much closer in time to the primary event in the cluster, indicating that the secondary events are likely triggered by static stress changes from the primary events.

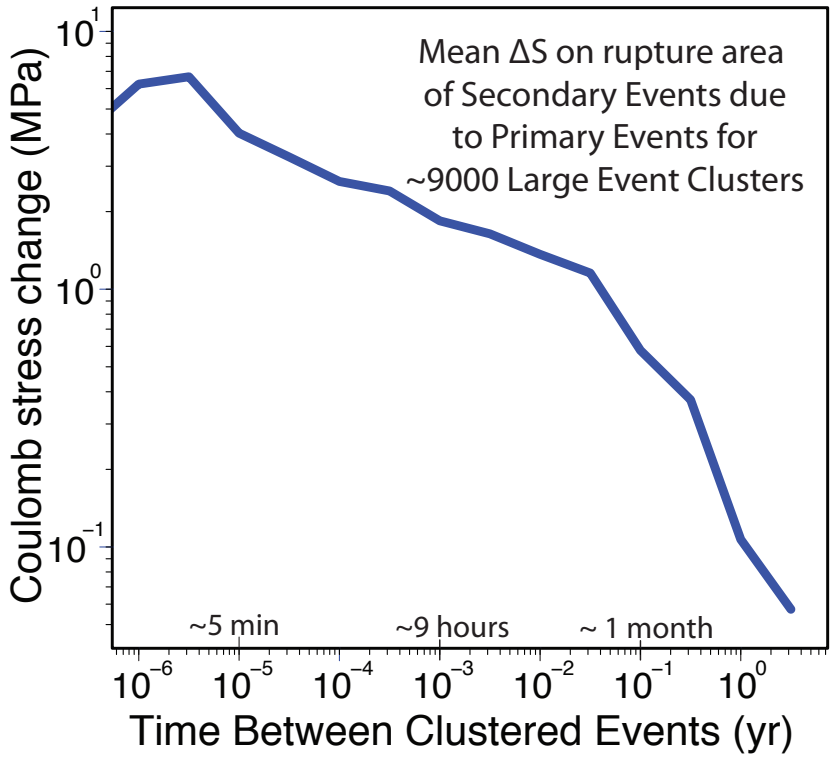

Figure 5.6: Coulomb stress change imparted on the rupture area of the secondary event in large-event clusters with the inter-event time between the primary and secondary events in the clusters. Data averaged over more than 9000 large-event clusters. In large event-clusters, when primary events transfer more stress onto the rupture area of the future secondary events, the secondary events occur closer in time to the primary events.

However, the stress transfer from primary clustered events to secondary events is not the only factor in the occurrence of large-event clusters. In addition to stress transfer, we also examined the role of prestress as a likely factor controlling the occurrence of large-event clusters. Specifically, as seen in Chapter 4, the elevated aftershock productivity in the interval between large earthquakes in large-event clusters suggests high pre-stresses on the rupture areas of the secondary event in the cluster.

This effect is illustrated by the example in Figure 5.1, where the primary event of the cluster and the independent event are nearly identical. The independent event

nucleated in nearly the same place (on the fault element immediately adjacent to the nucleation point of the primary clustered event) and the average slip in both mainshocks was nearly identical, but the independent event clearly did not trigger another large event. The independent event also ruptured in the same direction as the primary clustered event and stopped in the same location. More importantly, the independent event transferred the same amount of stress onto the next section of the fault as the primary clustered event, but that section of the fault did not experience another M≥7 event within 4 years. As the stress perturbation is the same for these clustered and non-clustered events, this highly productive aftershock sequence must also be related to the prestress on the nearby faults.

Figure 5.7 shows the prestress on the rupture area of the secondary event shown in Figure 5.1, for both the large-event cluster and the independent event. The color represents the absolute Coulomb stress on the rupture area of the secondary event in that large-event cluster, immediately prior to the primary event in the cluster (top) and immediately prior to the independent event (bottom and outlined in black). This is not the stress change, but the prestress on the section of the fault where the secondary event occurred. Over this area the absolute prestress is at least 2 MPa higher for nearly every fault element that ruptured in the secondary event and the average Coulomb stress over this area is also more than 2 MPa higher than before the independent event. Therefore, the stress states of faults is likely a determining factor for the aftershock productivity of the clustered events, as well as the occurrence of large-event clusters. More analysis is required to establish the prevalence of higher prestresses prior to secondary clustered events and the importance of these prestresses on the occurrence of large-event clusters.

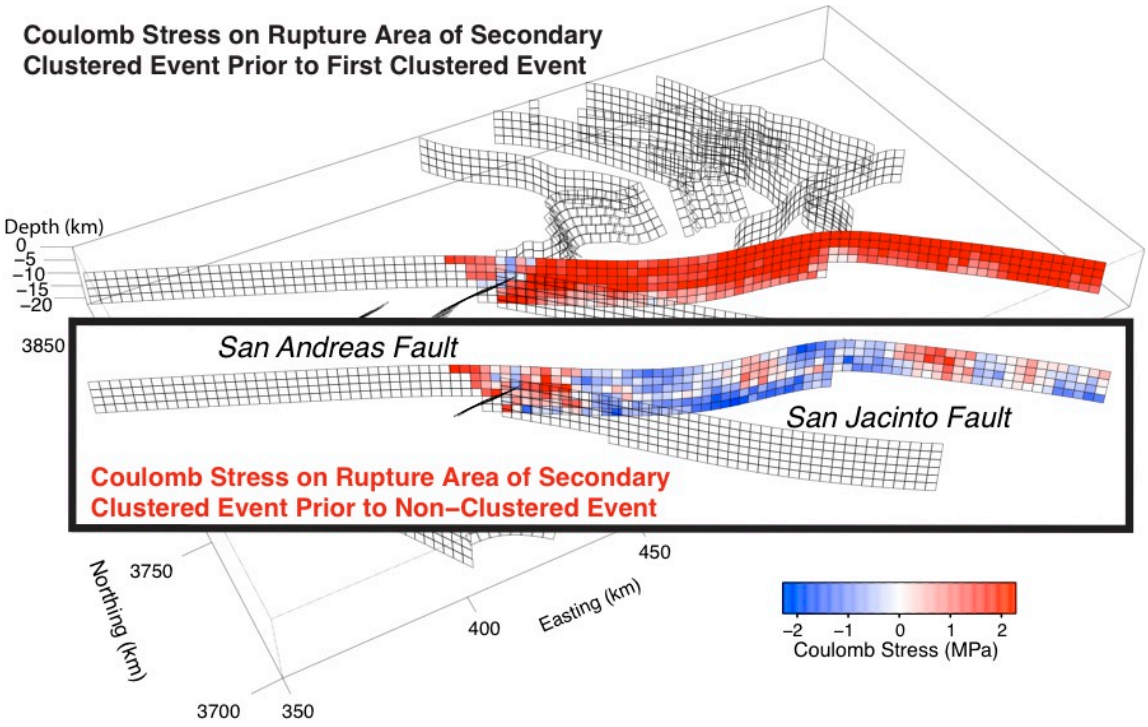

Figure 5.7: Comparison of the Coulomb stress on the rupture area of the secondary event in the large-event cluster shown in Figure 5.1 prior to the primary event of that cluster and that of the same fault elements prior to the independent (non-clustered) event shown in Figure 5.1, that was not followed by another event within 4 years. Even when the change in Coulomb stress is the same after a large event, the prestress on the rupture area of the secondary event is an important factor in the timing of that secondary event.

5.3.4 Clustering Probabilities Based on Aftershock Productivity

Chapter 4 presented clustering probabilities based solely on the number of clustered and independent events in the catalog. While such large-event probabilities can give valuable information about the earthquake hazard in an area, aftershock data appear to give significantly more predictive information. Figure 5.8 shows the probability of large-event clustering, i.e., the probability that an $M \geq 7$ event will be followed by another $M \geq 7$ event within 4 years, based on the number of aftershocks observed in one day following the mainshock. Observed clustering probabilities ($P_{Observed}$) were calculated

for primary events with different numbers of aftershocks within one day of the primary mainshock. All remaining probabilities are from the UCERF3 'Paleoseismic' Catalog, which had a rate- and state- *a*-value of 0.01. Dieterich (1994) demonstrated that the rate-state friction parameter *a* controls aftershock duration and hence aftershock productivity. In general, the rate of all events, and thus aftershocks, increases with increasing *a*-value. An *a*-value of 0.01 is in the middle of the range tested in this study (Table 4.1) of 0.08 to 0.12. These clustering probabilities for different numbers of aftershocks are fit by an exponential function for smoothing, which is plotted in Figure 5.8. For comparison, the red line in Figure 5.8 indicates the observed probability of clustering based solely on the general statistics of large-event clusters in the catalog. For example, from Figure 5.8, the overall probability of $M \geq 7$ clusters is on average just below 10%. However, the probability of additional $M \geq 7$ events is doubled if one were to observe ten $M \geq 4$ aftershocks within the first day following a mainshock.

Simulation results suggest that there is additional predictive information about impending large events beyond what current statistical methods for forecasting probabilities of large-event clusters provide. Figure 5.9 shows the probability that an additional $M \geq 7$ event will occur within 4 years after an $M \geq 7$ event for any given number of aftershocks with time from the mainshock. Similar to the probabilities in Figure 5.8, RSQSim probabilities (top) are calculated based on the observed number of large-event clusters with some number of aftershocks that occur after the first event of each large-event cluster. This is done for all inter-event times of the clustered events, and once again, an exponential function is fit to the data for smoothing. Reasenberg-Jones

probabilities (Figure 5.8, bottom) are based on the empirical Omori aftershock and Gutenberg-Richter magnitude frequency laws (Reasenberg and Jones, 1989). For example, if a simulated event has 10 aftershocks in 9 hours, there is a 40% to 50% chance that that event will be followed by another M≥7 event within 4 years, but if one were to only consider probabilities implied by empirical earthquake statistics, the probability would be assumed to be significantly lower, i.e. only 10% to 20%. The larger RSQSim probabilities suggest that high aftershock rates before large-event clusters seen in RSQSim indicate near-critical stresses for failure in large earthquakes. Although the stresses are known for every fault element in the RSQSim model at all times, this is impossible to know in the real world. As shown above, aftershocks occurring on the modeled faults, which might fail in another large earthquake, might serve as a proxy for the stress state on the fault and can be observed in real time. However, this cannot yet be applied to real-world earthquakes because we are only considering the aftershocks that occur on the explicitly modeled faults, which is only a subset of the aftershocks expected in nature because many of those events occur on smaller faults that are not built into our model.

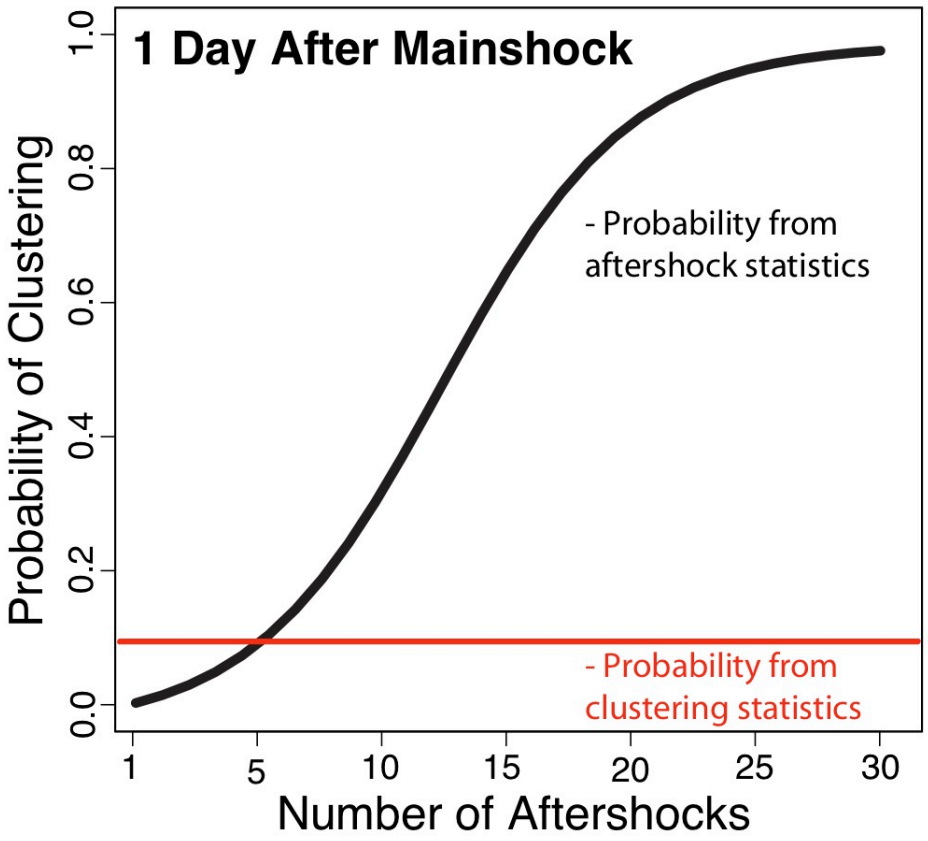

Figure 5.8: Probability of additional M≥7 event within 4 years, given the number of aftershocks in 1 day following a prior *M*≥ 7 event. Red line indicates the probability of additional *M*≥7 events based only on clustering statistics without considering the aftershock productivity. This indicates that recognizing highly productive aftershock sequences could improve short-term earthquake probabilities.

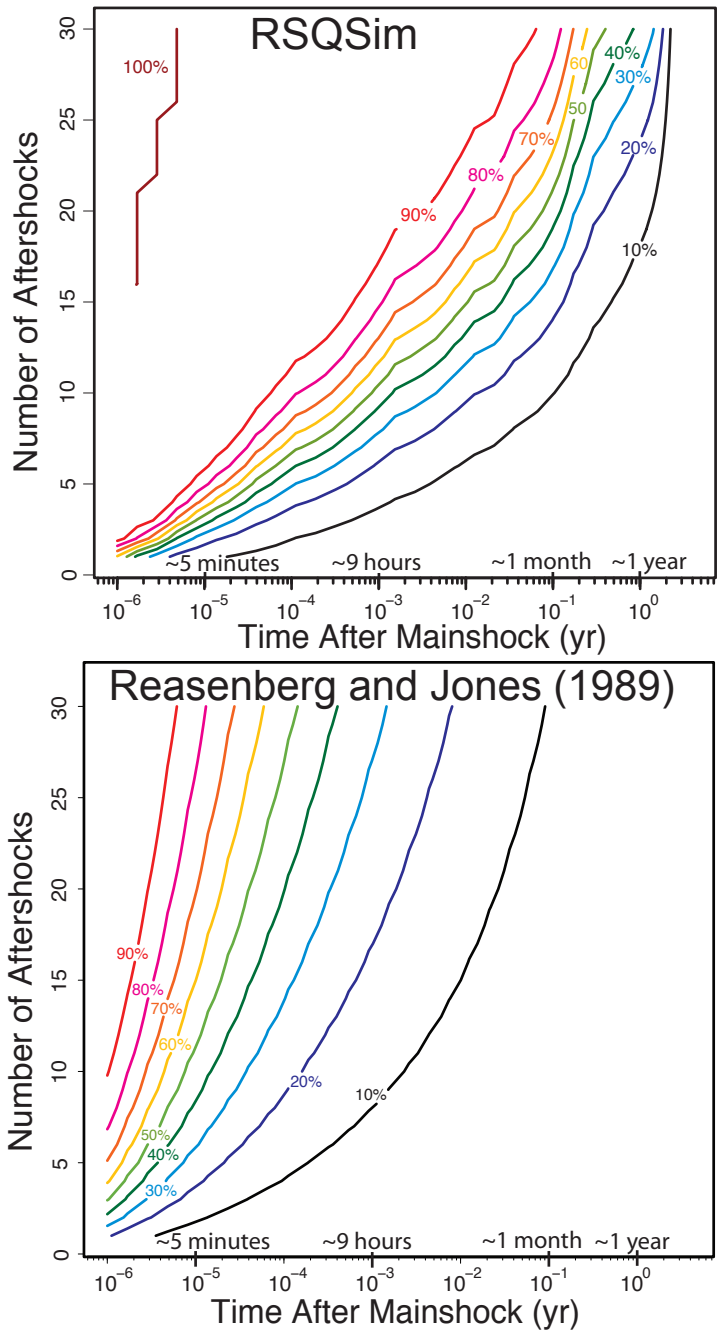

Figure 5.9: Colored contours indicate the probability that a second M≥7 event will occur within 4 years after an M≥7 event. RSQSim probabilities (left) are based on the number of aftershocks that occur after the first event of each large-event cluster. Reasenberg-Jones probabilities (right) are based on the empirical Omori aftershock and Gutenberg-Richter magnitude frequency laws. This indicates that aftershock statistics can give more predictive information about the occurrence of large event-clusters than probabilities based solely on empirical earthquake statistics.

5.4 Discussion and Conclusions

Earthquake scientists have searched for precursory signals for decades without any conclusive success. However, long-term simulated catalogs from RSQSim display several potential indicators that additional large earthquakes may shortly follow previous large earthquakes. Simulated events that become part of a large-event cluster have more productive aftershock sequences than large, independent events. Aftershocks of the primary events in large-event clusters correlate with the locations of secondary mainshocks, and migrate towards the secondary mainshock location with time leading up to nucleation of that event. RSQSim simulations also demonstrate accelerating aftershock sequences, in which aftershocks transition into foreshocks, and foreshocks of the secondary events tend to localize in the area of secondary mainshock hypocenters. The prestress on the faults where large-event clusters occur is also an important factor in the occurrence of large-event clusters and the aftershock productivity seems to indicate the areas that are already critically stressed and capable of producing additional large ($M \geq 7$) earthquakes. Additionally, simulation results suggest that there is additional predictive information about impending large events beyond what current statistical methods for forecasting probabilities of large-event clusters provide. The probability that a secondary large event will occur in RSQSim simulations, thus creating a large-event cluster, is significantly higher for mainshocks that have highly productive aftershocks sequences. This suggests that highly productive aftershock sequences could be used to forecast the occurrence of secondary events that become part of large-event clusters.

However, there are limitations to direct implementation for earthquake forecasts. The relatively small number of foreshocks observed for natural earthquakes may limit direct applicability of our observations of accelerating seismicity and localization of foreshocks. This study was able to identify patterns and associations with foreshocks and aftershocks because we had the ability to stack the foreshock/aftershock data for thousands of large events. The increased number of natural events at magnitudes smaller than we are able to simulate with a fault system this size may offset this problem, but that analysis is beyond the scope of this study. Additionally, our results for earthquake probability based on aftershocks are for the modeled faults only, which include only the major mapped faults. The vast majority of aftershocks in nature occur in the material surrounding those faults. We have the ability to add 'off-fault' seismicity to RSQSim catalogs, in which the stress changes due to large events are used to produce seismicity off of the modeled faults in areas with randomly oriented fault planes. However, this is only done after a long RSQSim simulation has completed and the 'off-fault' events do not feed back into the long RSQSim simulation. Future work will involve testing the addition of 'off-fault' seismicity and comparing the statistics to natural aftershock sequences and background seismicity rates.

5.5 References

Dieterich, J. H. (1988). Probability of earthquake recurrence with nonuniform stress rates and time-dependent failure. *Pure and Applied Geophysics*, *126*(2-4), 589-617.

Dieterich, J. H., and Richards-Dinger, K. B. (2010), Earthquake recurrence in simulated fault systems, Pure Appl. Geophysics., 167,1087–1104, doi:10.1007/s00024-010-0094-0.

Dewey, J. W. (1976). Seismicity of northern Anatolia: Seismological Society of America Bulletin, v. 66, p. 843–868.

Freed, A. M. (2005). Earthquake triggering by static, dynamic, and postseismic stress transfer, Annu. Rev. Earth Planet. Sci., 33, 335–367.

Geller, R. J. (1997). Earthquake Prediction: A Critical Review, Geophysics. J. Int, 131, 425-450.

Hardebeck, J. L., Nazareth, J. J., & Hauksson, E. (1998). The static stress change triggering model: Constraints from two southern California aftershock sequences. *Journal of Geophysical Research: Solid Earth (1978–2012)*,*103*(B10), 24427-24437.

Hudnut, K. W., Seeber, L., & Pacheco, J. (1989). Cross‐fault triggering in the November 1987 Superstition Hills Earthquake Sequence, Southern California.*Geophysical Research Letters*, *16*(2), 199-202.

Jackson, J., and McKenzie, D. (1988). The relationship between plate motions and seismic tremors, and the rates of active deformation in the Mediterranean and Middle East: Royal Astronomical Society Geophysical Journal, v. 93, p. 45–73.

King, G. C. P., Stein, R. S. and Lin, J. (1994). Static stress changes and the triggering of earthquakes, Bull. Seismol. Soc. Amer., 84, 935-953.

Reasenberg, P. A., & Jones, L. M. (1989). Earthquake hazard after a mainshock in California. *Science*, *243*(4895), 1173-1176.

Richter, C. F. (1958). Elementary Seismology, W. H. Freeman and Co., San Francisco, California.

Scholz, C. H., The Mechanics of Earthquakes and Faulting. Cambridge: Cambridge University Press, 1990.

Stein, R. S., Barka, A. A. and Dieterich, J. H. (1997). Progressive failure on the North Anatolian fault since 1939 by earthquake stress triggering, Geophysics Journal International, 128, 594–604.

Parsons, T., Stein, R. S., Simpson, R. W. and Reasenberg, P. A. (1999). Stress sensitivity of fault seismicity: A comparison between limited-offset oblique and major strike-slip faults, J. Geophysics. Res., 104, 20,183–20,202.

Chapter 6
Conclusions and Future Work

RSQSim has been successful at modeling several aspects of fault slip and earthquake occurrence and we believe that multi-event earthquake models with time-dependent nucleation based on rate-state friction, such as RSQSim, provide a viable physics-based method for modeling earthquake and fault system processes. These models not only provide a better understanding of earthquake hazard by improving our general knowledge of earthquake processes and statistics, but they can directly contribute to region-specific evaluation of long-term earthquake probabilities, which currently do not account for earthquake clustering. These models may also have an application for short-term operational earthquake forecasting which is, at present, primarily based on empirical earthquake statistics as a whole rather than the statistics of earthquakes in specific areas.

The first priority for continued work with RSQSim is to improve the California fault model and generate a California catalog that is both long enough to contribute relevant earthquake probabilities and statistics, as well as detailed enough to accurately describe all aspects of California seismicity. While the RSQSim California catalogs have already been shown to reproduce several aspects of known California seismicity including the statistics of foreshocks, aftershocks, and large-event clusters, as well as the statistics of inter-event times of smaller events, other aspects have not yet been included in a full California catalog. These include adding off-fault seismicity to improve the statistics and background rates for smaller events, specifically foreshocks and aftershocks

that occur off the large, mapped faults that are explicitly modeled. Some work has been done in an effort to compute off-fault seismicity for a completed RSQSim simulation, however the work is, as yet, incomplete and there is no computationally feasible way to have the off-fault seismicity interact with the long-term simulations. Additionally, including creeping sections in this fault model will also be completed as soon as possible. Viscoelasticity, in the form of deep creeping fault sections beneath the currently modeled faults is a top priority, as we believe it will increase the coefficient of variation of recurrence distributions and improve conditional probability calculations. Shallow creeping sections will also be added to the fault segments that are known to creep near the surface, such as part of the San Andreas Fault in Parkfield and the Hayward and Calaveras Faults further north. Furthermore, relaxing the segmentation of the UCERF fault models and modeling the major California faults with fully continuous faults may reduce any artificial boundaries or barriers to rupture that may artificially elevate clustering probabilities. Finally, the range of uncertain model parameters capable of satisfying the observation should be fully explored.

Results from this study can lso contribute directly to seismic hazard analysis and earthquake engineering efforts by providing source models for dynamic rupture simulations. Although dynamic rupture simulations can be very powerful tools, our results suggest that current methods may not be adequately estimating ground motions. Better initial stress distributions for dynamic models, in the form of evolved stresses from RSQSim simulations, will improve seismic hazard estimates and allow earthquake scientists to better evaluate the risk posed by earthquakes that occur in or near highly

populated areas. One application of the results of this project is to develop physics-based methods to define sets of initial stress conditions for use in generating suites of synthetic seismograms. These synthetic seismograms are instrumental in the design of critical structures such as bridges and power plants, in addition to schools, office buildings, and residences.

Suites of initial conditions for dynamic rupture models, as well as slip and stress drop values for kinematic ground motion calculations can be computed for any fault or fault system using RSQSim. Efforts are already underway to utilize CUBIT, a geometry and mesh generation toolkit from Sandia National Laboratories, to produce large, multi-fault meshes for FaultMod. RSQSim is able to simulate hundreds of thousands of events in large, complicated fault systems, so initial conditions can be generated and used to estimate ground motions for any modeled region or fault system. Specifically, models will be tailored to significant faults in California using long-term simulations with California fault models. Furthermore, the modeling process is not restricted to California fault systems. Any mapped fault system can be modeled for use in RSQSim simulations to generate earthquake probabilities and source models that will ultimately lead to the construction of safer, more stable structures. Additionally, long-term, multi-event RSQSim simulations can be used to calculate not just the fault conditions prior to large events but also the likelihood that specific large events will occur. These earthquake probabilities can be combined with source models to compute ground motion estimates for high probability events in fault system models.